GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 11:44:09 ; Search time 633.333 Seconds
(without alignments)
1984.655 Million cell updates/sec

Title: US-09-310-844C-23
Perfect score: 29
Sequence: 1 nnngauncuuunnguaagcccnangngnn 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 1733942

Minimum DB seq length: 0
Maximum DB seq length: 70

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	18	62.1	42	6	BD274270	BD274270 Identific
2	18	62.1	42	6	BD274271	BD274271 Identific
3	18	62.1	42	6	BD274272	BD274272 Identific
4	18	62.1	42	6	BD274273	BD274273 Identific
5	18	62.1	42	6	BD274275	BD274275 Identific
6	18	62.1	42	6	BD274278	BD274278 Identific
7	18	62.1	42	6	BD274279	BD274279 Identific
8	18	62.1	42	6	BD274280	BD274280 Identific
9	18	62.1	42	6	BD274281	BD274281 Identific
10	18	62.1	42	6	BD274283	BD274283 Identific
11	18	62.1	42	6	BD274284	BD274284 Identific
12	18	62.1	44	6	BD274277	BD274277 Identific
13	18	62.1	46	6	BD274238	BD274238 Identific
14	18	62.1	46	6	BD274240	BD274240 Identific
15	18	62.1	46	6	BD274241	BD274241 Identific
16	18	62.1	46	6	BD274242	BD274242 Identific
17	18	62.1	46	6	BD274243	BD274243 Identific
18	18	62.1	46	6	BD274246	BD274246 Identific
19	18	62.1	46	6	BD274247	BD274247 Identific
20	18	62.1	46	6	BD274248	BD274248 Identific
21	18	62.1	46	6	BD274249	BD274249 Identific
22	18	62.1	46	6	BD274252	BD274252 Identific
23	18	62.1	46	6	BD274253	BD274253 Identific
24	18	62.1	46	6	BD274256	BD274256 Identific
25	18	62.1	46	6	BD274257	BD274257 Identific
26	18	62.1	46	6	BD274258	BD274258 Identific
27	18	62.1	46	6	BD274259	BD274259 Identific
28	18	62.1	46	6	BD274260	BD274260 Identific
29	18	62.1	46	6	BD274262	BD274262 Identific
30	18	62.1	46	6	BD274263	BD274263 Identific
31	18	62.1	46	6	BD274264	BD274264 Identific
32	18	62.1	46	6	BD274265	BD274265 Identific
33	18	62.1	46	6	BD274268	BD274268 Identific
34	18	62.1	46	6	BD274269	BD274269 Identific
35	15	51.7	46	6	BD274237	BD274237 Identific
36	15	51.7	46	6	BD274251	BD274251 Identific
37	15	51.7	46	6	BD274255	BD274255 Identific
38	15	51.7	46	6	BD274267	BD274267 Identific
39	14.8	51.0	42	6	BD274276	BD274276 Identific
40	13	44.8	42	6	BD274274	BD274274 Identific
41	13	44.8	42	6	BD274282	BD274282 Identific
c 42	12.8	44.1	53	10	MDTRVNJB	X63580 M.domesticu
43	12.8	44.1	57	8	AJ591789	AJ591789 Arabidops
44	12.8	44.1	68	6	AR356644	AR356644 Sequence
45	12.2	42.1	27	6	AR364272	AR364272 Sequence

ALIGNMENTS

RESULT 1
BD274270
LOCUS BD274270 42 bp DNA linear PAT 17-JUL-2003
DEFINITION Identification of molecular interaction sites in RNA for novel drug discovery.
ACCESSION BD274270
VERSION BD274270.1 GI:33084038
KEYWORDS JP 2002526030-A/237.
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 42)
AUTHORS Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
TITLE Identification of molecular interaction sites in RNA for novel drug discovery

```
  JOURNAL   Patent: JP 2002526030-A 237 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/237
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .42
            FT                   /organism='Artificial Sequence'.
FEATURES              Location/Qualifiers
     source           1. .42
                      /organism="synthetic construct"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:32630"
ORIGIN

  Query Match             62.1%;  Score 18;  DB 6;  Length 42;
  Best Local Similarity   54.2%;  Pred. No. 5.6;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||: |::: |:|||||| | | |
Db          7 GATTCTTTTTGTAAGCCCTACGGG 30


RESULT 2
BD274271
LOCUS       BD274271                  42 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274271
VERSION     BD274271.1  GI:33084039
KEYWORDS    JP 2002526030-A/238.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 238 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/238
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .42
            FT                   /organism='Artificial Sequence'.
FEATURES              Location/Qualifiers
     source           1. .42
                      /organism="synthetic construct"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:32630"
ORIGIN

  Query Match             62.1%;  Score 18;  DB 6;  Length 42;
  Best Local Similarity   54.2%;  Pred. No. 5.6;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||: |::: |:|||||| | | |
Db          7 GATTCTTTTTGTAAGCCCTAGGCG 30


RESULT 3
BD274272
LOCUS       BD274272                  42 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274272
VERSION     BD274272.1  GI:33084040
KEYWORDS    JP 2002526030-A/239.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 239 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/239
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .42
            FT                   /organism='Artificial Sequence'.
FEATURES              Location/Qualifiers
     source           1. .42
                      /organism="synthetic construct"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:32630"
ORIGIN

  Query Match             62.1%;  Score 18;  DB 6;  Length 42;
  Best Local Similarity   54.2%;  Pred. No. 5.6;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||: |::: |:|||||| | | |
Db          7 GATTCTTTTTGTAAGCCCCAAGGG 30


RESULT 4
BD274273
LOCUS       BD274273                  42 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274273
VERSION     BD274273.1  GI:33084041
KEYWORDS    JP 2002526030-A/240.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 240 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/240
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .42
            FT                   /organism='Artificial Sequence'.
FEATURES              Location/Qualifiers
```

```
     source          1. .42
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             62.1%;  Score 18;  DB 6;  Length 42;
  Best Local Similarity   54.2%;  Pred. No. 5.6;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||: |::: |:|||||| | | |
Db          7 GATTCTTTTTGTAAGCCCCAAGGG 30


RESULT 5
BD274275
LOCUS       BD274275                  42 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274275
VERSION     BD274275.1  GI:33084043
KEYWORDS    JP 2002526030-A/242.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 242 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/242
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .42
            FT                   /organism='Artificial Sequence'.
FEATURES              Location/Qualifiers
     source           1. .42
                      /organism="synthetic construct"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:32630"
ORIGIN

  Query Match             62.1%;  Score 18;  DB 6;  Length 42;
  Best Local Similarity   54.2%;  Pred. No. 5.6;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||: |::: |:|||||| | | |
Db          7 GATTCTTTTTGTAAGCCCTAGGGG 30


RESULT 6
BD274278
LOCUS       BD274278                  42 bp    RNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274278
VERSION     BD274278.1  GI:33084046
KEYWORDS    JP 2002526030-A/245.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 245 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/245
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .42
            FT                   /organism='Artificial Sequence'.
FEATURES              Location/Qualifiers
     source           1. .42
                      /organism="synthetic construct"
                      /mol_type="genomic RNA"
                      /db_xref="taxon:32630"
ORIGIN

  Query Match             62.1%;  Score 18;  DB 6;  Length 42;
  Best Local Similarity   54.2%;  Pred. No. 5.6;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||: |::: |:|||||| | | |
Db          7 GATTCTTTTTGTAAGCCCTACGGG 30


RESULT 7
BD274279
LOCUS       BD274279                  42 bp    RNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274279
VERSION     BD274279.1  GI:33084047
KEYWORDS    JP 2002526030-A/246.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 246 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/246
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .42
            FT                   /organism='Artificial Sequence'.
FEATURES              Location/Qualifiers
     source           1. .42
                      /organism="synthetic construct"
                      /mol_type="genomic RNA"
                      /db_xref="taxon:32630"
ORIGIN

  Query Match             62.1%;  Score 18;  DB 6;  Length 42;
  Best Local Similarity   54.2%;  Pred. No. 5.6;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||: |::: |:|||||| | | |
Db          7 GATTCTTTTTGTAAGCCCTAGGCG 30
```

```
RESULT 8
BD274280
LOCUS       BD274280                  42 bp    RNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274280
VERSION     BD274280.1  GI:33084048
KEYWORDS    JP 2002526030-A/247.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 247 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/247
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .42
            FT                   /organism='Artificial Sequence'.
FEATURES              Location/Qualifiers
     source           1. .42
                      /organism="synthetic construct"
                      /mol_type="genomic RNA"
                      /db_xref="taxon:32630"
ORIGIN

  Query Match           62.1%;  Score 18;  DB 6;  Length 42;
  Best Local Similarity   54.2%;  Pred. No. 5.6;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||: |::: |:|||||| | | |
Db          7 GATTCTTTTTGTAAGCCCCAAGGG 30


RESULT 9
BD274281
LOCUS       BD274281                  42 bp    RNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274281
VERSION     BD274281.1  GI:33084049
KEYWORDS    JP 2002526030-A/248.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 248 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/248
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .42
            FT                   /organism='Artificial Sequence'.
FEATURES              Location/Qualifiers
     source           1. .42
                      /organism="synthetic construct"
                      /mol_type="genomic RNA"
                      /db_xref="taxon:32630"
ORIGIN

  Query Match           62.1%;  Score 18;  DB 6;  Length 42;
  Best Local Similarity   54.2%;  Pred. No. 5.6;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||: |::: |:|||||| | | |
Db          7 GATTCTTTTTGTAAGCCCTAGGGG 30


RESULT 10
BD274283
LOCUS       BD274283                  42 bp    RNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274283
VERSION     BD274283.1  GI:33084051
KEYWORDS    JP 2002526030-A/250.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 250 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/250
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .42
            FT                   /organism='Artificial Sequence'.
FEATURES              Location/Qualifiers
     source           1. .42
                      /organism="synthetic construct"
                      /mol_type="genomic RNA"
                      /db_xref="taxon:32630"
ORIGIN

  Query Match           62.1%;  Score 18;  DB 6;  Length 42;
  Best Local Similarity   54.2%;  Pred. No. 5.6;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||: |::: |:|||||| | | |
Db          7 GATTCTTTTTGTAAGCCCTAGGGG 30


RESULT 11
BD274284
LOCUS       BD274284                  42 bp    RNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274284
VERSION     BD274284.1  GI:33084052
KEYWORDS    JP 2002526030-A/251.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
```

```
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 251 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/251
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .42
            FT                   /organism='Artificial Sequence'.
FEATURES              Location/Qualifiers
     source           1. .42
                      /organism="synthetic construct"
                      /mol_type="genomic RNA"
                      /db_xref="taxon:32630"
ORIGIN

  Query Match            62.1%;  Score 18;  DB 6;  Length 42;
  Best Local Similarity  54.2%;  Pred. No. 5.6;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||: |:::  |:|||||| | | |
Db          7 GATCCTTTCTGTAAGCCCTACGGG 30


RESULT 12
BD274277
LOCUS       BD274277                  44 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274277
VERSION     BD274277.1  GI:33084045
KEYWORDS    JP 2002526030-A/244.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 44)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 244 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/244
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence CC   N is any nucleotide
            CC   N is any nucleotide
            CC   N is any nucleotide
            CC   N is any nucleotide
            CC   N is any nucleotide
            CC   N is any nucleotide
            CC   N is any nucleotide
            CC   N is any nucleotide
            FH   Key             Location/Qualifiers
            FT   misc_feature    (1). .(2)
            FT   misc_feature    (4). .(7)
            FT   misc_feature    (11)
            FT   misc_feature    (16). .(17)
            FT   misc_feature    (26)
            FT   misc_feature    (28)
            FT   misc_feature    (30)
            FT   misc_feature    (32). .(35)
            FT   misc_feature    (39). .(44).
FEATURES              Location/Qualifiers
     source           1. .44
                      /organism="synthetic construct"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:32630"
ORIGIN

  Query Match            62.1%;  Score 18;  DB 6;  Length 44;
  Best Local Similarity  79.2%;  Pred. No. 5.6;
  Matches   19;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||:||:::|||:||||||||||||
Db          8 GATNCTTTNNGTAAGCCCNANGNG 31


RESULT 13
BD274238
LOCUS       BD274238                  46 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274238
VERSION     BD274238.1  GI:33084006
KEYWORDS    JP 2002526030-A/205.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 205 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/205
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .46
            FT                   /organism='Artificial Sequence'.
FEATURES              Location/Qualifiers
     source           1. .46
                      /organism="synthetic construct"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:32630"
ORIGIN

  Query Match            62.1%;  Score 18;  DB 6;  Length 46;
  Best Local Similarity  54.2%;  Pred. No. 5.6;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||: |:::  |:|||||| | | |
Db         22 GATTCTTTTTGTAAGCCCTACGGG 45


RESULT 14
BD274240
LOCUS       BD274240                  46 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274240
VERSION     BD274240.1  GI:33084008
KEYWORDS    JP 2002526030-A/207.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 46)
```

```
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 207 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/207
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .46
            FT                   /organism='Artificial Sequence'.
FEATURES              Location/Qualifiers
     source           1. .46
                      /organism="synthetic construct"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:32630"
ORIGIN

  Query Match            62.1%;  Score 18;  DB 6;  Length 46;
  Best Local Similarity  54.2%;  Pred. No. 5.6;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||: |:::  |:|||||| | | |
Db         22 GATTCTTTTTGTAAGCCCTAGGGG 45


RESULT 15
BD274241
LOCUS       BD274241                 46 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274241
VERSION     BD274241.1  GI:33084009
KEYWORDS    JP 2002526030-A/208.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 208 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/208
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .46
            FT                   /organism='Artificial Sequence'.
FEATURES              Location/Qualifiers
     source           1. .46
                      /organism="synthetic construct"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:32630"
ORIGIN

  Query Match            62.1%;  Score 18;  DB 6;  Length 46;
  Best Local Similarity  54.2%;  Pred. No. 5.6;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||: |:::  |:|||||| | | |
Db         22 GATTCTTTTTGTAAGCCCCAAGGG 45

Search completed: March 23, 2004, 15:25:07
Job time : 634.333 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 11:34:35 ; Search time 235.333 Seconds
(without alignments)
523.503 Million cell updates/sec

Title: US-09-310-844C-23
Perfect score: 29
Sequence: 1 nnngauncuuunnguaagcccnangngnn 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 3369620

Minimum DB seq length: 0
Maximum DB seq length: 70

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	18	62.1	29	3	AAA70827	Aaa70827 Molecular
2	18	62.1	29	3	AAA70829	Aaa70829 Molecular
3	18	62.1	29	3	AAA70828	Aaa70828 Molecular
4	18	62.1	29	3	AAA70830	Aaa70830 Molecular
5	18	62.1	42	3	AAA71121	Aaa71121 Molecular
6	18	62.1	42	3	AAA71128	Aaa71128 Molecular
7	18	62.1	42	3	AAA71123	Aaa71123 Molecular
8	18	62.1	42	3	AAA71113	Aaa71113 Molecular
9	18	62.1	42	3	AAA71124	Aaa71124 Molecular
10	18	62.1	42	3	AAA71132	Aaa71132 Molecular
11	18	62.1	42	3	AAA71120	Aaa71120 Molecular
12	18	62.1	42	3	AAA71114	Aaa71114 Molecular
13	18	62.1	42	3	AAA71118	Aaa71118 Molecular
14	18	62.1	42	3	AAA71119	Aaa71119 Molecular
15	18	62.1	42	3	AAA71126	Aaa71126 Molecular
16	18	62.1	42	3	AAA71131	Aaa71131 Molecular
17	18	62.1	42	3	AAA71127	Aaa71127 Molecular
18	18	62.1	42	3	AAA71116	Aaa71116 Molecular
19	18	62.1	42	3	AAA71115	Aaa71115 Molecular
20	18	62.1	42	3	AAA71129	Aaa71129 Molecular
21	18	62.1	44	3	ABK87476	Abk87476 Interleuk
22	18	62.1	44	3	AAA71112	Aaa71112 Molecular
23	18	62.1	44	3	AAA71125	Aaa71125 Molecular
24	18	62.1	44	3	AAA71133	Aaa71133 Molecular
25	18	62.1	45	3	AAA70826	Aaa70826 Molecular
26	18	62.1	45	3	AAA70825	Aaa70825 Molecular
27	18	62.1	45	3	AAA70824	Aaa70824 Molecular
28	18	62.1	46	3	AAA71089	Aaa71089 Molecular
29	18	62.1	46	3	AAA71094	Aaa71094 Molecular
30	18	62.1	46	3	AAA71085	Aaa71085 Molecular
31	18	62.1	46	3	AAA71110	Aaa71110 Molecular
32	18	62.1	46	3	AAA71111	Aaa71111 Molecular
33	18	62.1	46	3	AAA71095	Aaa71095 Molecular
34	18	62.1	46	3	AAA71106	Aaa71106 Molecular
35	18	62.1	46	3	AAA71107	Aaa71107 Molecular
36	18	62.1	46	3	AAA71109	Aaa71109 Molecular
37	18	62.1	46	3	AAA71087	Aaa71087 Molecular
38	18	62.1	46	3	AAA71093	Aaa71093 Molecular
39	18	62.1	46	3	AAA71096	Aaa71096 Molecular
40	18	62.1	46	3	AAA71099	Aaa71099 Molecular
41	18	62.1	46	3	AAA71088	Aaa71088 Molecular
42	18	62.1	46	3	AAA71100	Aaa71100 Molecular
43	18	62.1	46	3	AAA71105	Aaa71105 Molecular
44	18	62.1	46	3	AAA71103	Aaa71103 Molecular
45	18	62.1	46	3	AAA71090	Aaa71090 Molecular

ALIGNMENTS

RESULT 1
AAA70827
ID AAA70827 standard; RNA; 29 BP.
XX
AC AAA70827;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #27.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Synthetic.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Claim 235; Page 235; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The

CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 29 BP; 4 A; 4 C; 5 G; 0 T; 5 U; 11 Other;

Query Match 62.1%; Score 18; DB 3; Length 29;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GAUNCUUUNNGUAAGCCCNANGNG 27
||||||||||||||||||||||||
Db 4 GAUNCUUUNNGUAAGCCCNANGNG 27

RESULT 2
AAA70829
ID AAA70829 standard; RNA; 29 BP.
XX
AC AAA70829;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #29.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Mus sp.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Claim 235; Page 235; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 29 BP; 8 A; 6 C; 6 G; 0 T; 9 U; 0 Other;

Query Match 62.1%; Score 18; DB 3; Length 29;
Best Local Similarity 75.0%; Pred. No. 2.1;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 GAUNCUUUNNGUAAGCCCNANGNG 27
||| |||| |||||||| | | |
Db 4 GAUUCUUUUUGUAAGCCCCAAGGG 27

RESULT 3
AAA70828
ID AAA70828 standard; RNA; 29 BP.
XX
AC AAA70828;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #28.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Homo sapiens.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Claim 235; Page 235; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary

CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 29 BP; 5 A; 5 C; 7 G; 0 T; 12 U; 0 Other;

Query Match 62.1%; Score 18; DB 3; Length 29;
Best Local Similarity 75.0%; Pred. No. 2.1;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 GAUNCUUUNNGUAAGCCCNANGNG 27
||| |||| |||||||| | | |
Db 4 GAUUCUUUUUGUAAGCCCUAGGGG 27

RESULT 4
AAA70830
ID AAA70830 standard; RNA; 29 BP.
XX
AC AAA70830;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #30.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Rattus sp.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Claim 235; Page 235; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 29 BP; 8 A; 6 C; 6 G; 0 T; 9 U; 0 Other;

Query Match 62.1%; Score 18; DB 3; Length 29;
Best Local Similarity 75.0%; Pred. No. 2.1;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 GAUNCUUUNNGUAAGCCCNANGNG 27
||| |||| |||||||| | | |
Db 4 GAUUCUUUUUGUAAGCCCCAAGGG 27

RESULT 5
AAA71121
ID AAA71121 standard; DNA; 42 BP.
XX
AC AAA71121;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #127.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 125; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an

CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 42 BP; 13 A; 7 C; 7 G; 15 T; 0 U; 0 Other;

Query Match 62.1%; Score 18; DB 3; Length 42;
Best Local Similarity 54.2%; Pred. No. 2.2;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 4 GAUNCUUUNNGUAAGCCCNANGNG 27
||: |::: |:|||||| | | |
Db 7 GATTCTTTTTGTAAGCCCCAAGGG 30

RESULT 6
AAA71128
ID AAA71128 standard; RNA; 42 BP.
XX
AC AAA71128;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #197.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 126; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 42 BP; 13 A; 7 C; 7 G; 0 T; 15 U; 0 Other;

Query Match 62.1%; Score 18; DB 3; Length 42;
Best Local Similarity 75.0%; Pred. No. 2.2;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 GAUNCUUUNNGUAAGCCCNANGNG 27
||| |||| |||||||| | | |
Db 7 GAUUCUUUUUGUAAGCCCCAAGGG 30

RESULT 7
AAA71123
ID AAA71123 standard; DNA; 42 BP.
XX
AC AAA71123;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #129.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 125; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4

CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 42 BP; 9 A; 6 C; 9 G; 18 T; 0 U; 0 Other;

Query Match 62.1%; Score 18; DB 3; Length 42;
Best Local Similarity 54.2%; Pred. No. 2.2;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 4 GAUNCUUUNNGUAAGCCCNANGNG 27
||: |::: |:|||||| | | |
Db 7 GATTCTTTTTGTAAGCCCTAGGGG 30

RESULT 8
AAA71113
ID AAA71113 standard; RNA; 42 BP.
XX
AC AAA71113;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #189.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 122; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 42 BP; 12 A; 7 C; 6 G; 0 T; 17 U; 0 Other;

Query Match 62.1%; Score 18; DB 3; Length 42;
Best Local Similarity 75.0%; Pred. No. 2.2;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 GAUNCUUUNNGUAAGCCCNANGNG 27
||| |||| |||||||| | | |
Db 7 GAUUCUUUUUGUAAGCCCUACGGG 30

RESULT 9
AAA71124
ID AAA71124 standard; DNA; 42 BP.
XX
AC AAA71124;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #130.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 125; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3

```
CC   nucleotides forming a second side of the first ds region; (2) a purified
CC   and isolated RNA fragment comprising the human sequence
CC   UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC   used for identifying agents which modulate the activity of biomolecules,
CC   particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC   or industrial compounds
XX
SQ   Sequence 42 BP; 11 A; 10 C; 7 G; 14 T; 0 U; 0 Other;

  Query Match            62.1%;  Score 18;  DB 3;  Length 42;
  Best Local Similarity  54.2%;  Pred. No. 2.2;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||: |:::  |:|||||| | | |
Db          7 GATCCTTTCTGTAAGCCCTACGGG 30


RESULT 10
AAA71132
ID   AAA71132 standard; RNA; 42 BP.
XX
AC   AAA71132;
XX
DT   27-APR-2001  (first entry)
XX
DE   Molecular interaction site RNA #201.
XX
KW   Modulator; identification; molecular interaction; virtual library; ss.
XX
OS   Unidentified.
XX
PN   WO9958947-A2.
XX
PD   18-NOV-1999.
XX
PF   12-MAY-1999;   99WO-US010361.
XX
PR   12-MAY-1998;   98US-00076404.
PR   12-MAY-1998;   98US-0085092P.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Ecker DJ,  Griffey R,  Crooke ST,  Sampath R,  Swayze E,  Mohan V;
PI   Hofstadler S,  Mcneil J;
XX
DR   WPI; 2000-086439/07.
XX
PT   Identifying compounds which modulate activity of target biomolecules,
PT   used to provide compounds which can be used as pharmacological,
PT   agricultural and industrial compounds.
XX
PS   Example 7; Fig 126; 405pp; English.
XX
CC   This invention describes a novel method for identifying compounds which
CC   modulate the activity of a target biomolecule. The method uses 3-
CC   dimensional representations of the biomolecule and a library of compounds
CC   and comprises (a) identifying at least one molecular interaction site of
CC   the target RNA; (b) generating in silico a virtual library of compounds
CC   predicted or calculated to interact with the molecular interaction site;
CC   and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC   with members of the virtual library of compounds to generate a hierarchy
CC   of the compounds ranked in accordance with their respective ability to
CC   form physical interactions with the molecular interaction site. The
CC   method also describes (1) RNA comprising a joined sequence of at least 24
CC   nucleotides but not more than 70 nucleotides and having secondary
CC   structure defined by: (a) 3 nucleotides forming a first side of a first
CC   double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC   internal loop region; (c) 4 nucleotides forming a first side of a second
CC   ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC   nucleotides forming a second side of the second ds region; (f) 4
CC   nucleotides forming a second side of the internal loop region; and (g) 3
CC   nucleotides forming a second side of the first ds region; (2) a purified
CC   and isolated RNA fragment comprising the human sequence
CC   UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC   used for identifying agents which modulate the activity of biomolecules,
CC   particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC   or industrial compounds
XX
SQ   Sequence 42 BP; 11 A; 10 C; 7 G; 0 T; 14 U; 0 Other;

  Query Match            62.1%;  Score 18;  DB 3;  Length 42;
  Best Local Similarity  75.0%;  Pred. No. 2.2;
  Matches   18;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||| ||||  ||||||||| | | |
Db          7 GAUCCUUUCUGUAAGCCCUACGGG 30


RESULT 11
AAA71120
ID   AAA71120 standard; DNA; 42 BP.
XX
AC   AAA71120;
XX
DT   27-APR-2001  (first entry)
XX
DE   Molecular interaction site DNA #126.
XX
KW   Modulator; identification; molecular interaction; virtual library; ss.
XX
OS   Unidentified.
XX
PN   WO9958947-A2.
XX
PD   18-NOV-1999.
XX
PF   12-MAY-1999;   99WO-US010361.
XX
PR   12-MAY-1998;   98US-00076404.
PR   12-MAY-1998;   98US-0085092P.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Ecker DJ,  Griffey R,  Crooke ST,  Sampath R,  Swayze E,  Mohan V;
PI   Hofstadler S,  Mcneil J;
XX
DR   WPI; 2000-086439/07.
XX
PT   Identifying compounds which modulate activity of target biomolecules,
PT   used to provide compounds which can be used as pharmacological,
PT   agricultural and industrial compounds.
XX
PS   Example 7; Fig 125; 405pp; English.
XX
CC   This invention describes a novel method for identifying compounds which
CC   modulate the activity of a target biomolecule. The method uses 3-
CC   dimensional representations of the biomolecule and a library of compounds
CC   and comprises (a) identifying at least one molecular interaction site of
CC   the target RNA; (b) generating in silico a virtual library of compounds
CC   predicted or calculated to interact with the molecular interaction site;
CC   and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC   with members of the virtual library of compounds to generate a hierarchy
CC   of the compounds ranked in accordance with their respective ability to
CC   form physical interactions with the molecular interaction site. The
CC   method also describes (1) RNA comprising a joined sequence of at least 24
CC   nucleotides but not more than 70 nucleotides and having secondary
CC   structure defined by: (a) 3 nucleotides forming a first side of a first
CC   double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC   internal loop region; (c) 4 nucleotides forming a first side of a second
CC   ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC   nucleotides forming a second side of the second ds region; (f) 4
CC   nucleotides forming a second side of the internal loop region; and (g) 3
CC   nucleotides forming a second side of the first ds region; (2) a purified
CC   and isolated RNA fragment comprising the human sequence
```

```
CC   UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC   used for identifying agents which modulate the activity of biomolecules,
CC   particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC   or industrial compounds
XX
SQ   Sequence 42 BP; 13 A; 7 C; 7 G; 15 T; 0 U; 0 Other;

  Query Match             62.1%;  Score 18;  DB 3;  Length 42;
  Best Local Similarity   54.2%;  Pred. No. 2.2;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||: |::: |:|||||| | | |
Db          7 GATTCTTTTTGTAAGCCCCAAGGG 30


RESULT 12
AAA71114
ID   AAA71114 standard; RNA; 42 BP.
XX
AC   AAA71114;
XX
DT   27-APR-2001  (first entry)
XX
DE   Molecular interaction site RNA #190.
XX
KW   Modulator; identification; molecular interaction; virtual library; ss.
XX
OS   Unidentified.
XX
PN   WO9958947-A2.
XX
PD   18-NOV-1999.
XX
PF   12-MAY-1999;   99WO-US010361.
XX
PR   12-MAY-1998;   98US-00076404.
PR   12-MAY-1998;   98US-0085092P.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Ecker DJ,  Griffey R,  Crooke ST,  Sampath R,  Swayze E,  Mohan V;
PI   Hofstadler S,  Mcneil J;
XX
DR   WPI; 2000-086439/07.
XX
PT   Identifying compounds which modulate activity of target biomolecules,
PT   used to provide compounds which can be used as pharmacological,
PT   agricultural and industrial compounds.
XX
PS   Example 7; Fig 122; 405pp; English.
XX
CC   This invention describes a novel method for identifying compounds which
CC   modulate the activity of a target biomolecule. The method uses 3-
CC   dimensional representations of the biomolecule and a library of compounds
CC   and comprises (a) identifying at least one molecular interaction site of
CC   the target RNA; (b) generating in silico a virtual library of compounds
CC   predicted or calculated to interact with the molecular interaction site;
CC   and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC   with members of the virtual library of compounds to generate a hierarchy
CC   of the compounds ranked in accordance with their respective ability to
CC   form physical interactions with the molecular interaction site. The
CC   method also describes (1) RNA comprising a joined sequence of at least 24
CC   nucleotides but not more than 70 nucleotides and having secondary
CC   structure defined by: (a) 3 nucleotides forming a first side of a first
CC   double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC   internal loop region; (c) 4 nucleotides forming a first side of a second
CC   ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC   nucleotides forming a second side of the second ds region; (f) 4
CC   nucleotides forming a second side of the internal loop region; and (g) 3
CC   nucleotides forming a second side of the first ds region; (2) a purified
CC   and isolated RNA fragment comprising the human sequence
CC   UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC   used for identifying agents which modulate the activity of biomolecules,
CC   particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC   or industrial compounds
XX
SQ   Sequence 42 BP; 11 A; 8 C; 7 G; 0 T; 16 U; 0 Other;

  Query Match             62.1%;  Score 18;  DB 3;  Length 42;
  Best Local Similarity   75.0%;  Pred. No. 2.2;
  Matches   18;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||| ||||  ||||||||| | | |
Db          7 GAUUCUUUUUGUAAGCCCUAGGCG 30


RESULT 13
AAA71118
ID   AAA71118 standard; DNA; 42 BP.
XX
AC   AAA71118;
XX
DT   27-APR-2001  (first entry)
XX
DE   Molecular interaction site DNA #124.
XX
KW   Modulator; identification; molecular interaction; virtual library; ss.
XX
OS   Unidentified.
XX
PN   WO9958947-A2.
XX
PD   18-NOV-1999.
XX
PF   12-MAY-1999;   99WO-US010361.
XX
PR   12-MAY-1998;   98US-00076404.
PR   12-MAY-1998;   98US-0085092P.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Ecker DJ,  Griffey R,  Crooke ST,  Sampath R,  Swayze E,  Mohan V;
PI   Hofstadler S,  Mcneil J;
XX
DR   WPI; 2000-086439/07.
XX
PT   Identifying compounds which modulate activity of target biomolecules,
PT   used to provide compounds which can be used as pharmacological,
PT   agricultural and industrial compounds.
XX
PS   Example 7; Fig 125; 405pp; English.
XX
CC   This invention describes a novel method for identifying compounds which
CC   modulate the activity of a target biomolecule. The method uses 3-
CC   dimensional representations of the biomolecule and a library of compounds
CC   and comprises (a) identifying at least one molecular interaction site of
CC   the target RNA; (b) generating in silico a virtual library of compounds
CC   predicted or calculated to interact with the molecular interaction site;
CC   and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC   with members of the virtual library of compounds to generate a hierarchy
CC   of the compounds ranked in accordance with their respective ability to
CC   form physical interactions with the molecular interaction site. The
CC   method also describes (1) RNA comprising a joined sequence of at least 24
CC   nucleotides but not more than 70 nucleotides and having secondary
CC   structure defined by: (a) 3 nucleotides forming a first side of a first
CC   double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC   internal loop region; (c) 4 nucleotides forming a first side of a second
CC   ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC   nucleotides forming a second side of the second ds region; (f) 4
CC   nucleotides forming a second side of the internal loop region; and (g) 3
CC   nucleotides forming a second side of the first ds region; (2) a purified
CC   and isolated RNA fragment comprising the human sequence
CC   UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC   used for identifying agents which modulate the activity of biomolecules,
```

CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 42 BP; 12 A; 7 C; 6 G; 17 T; 0 U; 0 Other;

Query Match 62.1%; Score 18; DB 3; Length 42;
Best Local Similarity 54.2%; Pred. No. 2.2;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 4 GAUNCUUUNNGUAAGCCCNANGNG 27
||: |::: |:|||||| | | |
Db 7 GATTCTTTTTGTAAGCCCTACGGG 30

RESULT 14
AAA71119
ID AAA71119 standard; DNA; 42 BP.
XX
AC AAA71119;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #125.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 125; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 42 BP; 11 A; 8 C; 7 G; 16 T; 0 U; 0 Other;

Query Match 62.1%; Score 18; DB 3; Length 42;
Best Local Similarity 54.2%; Pred. No. 2.2;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 4 GAUNCUUUNNGUAAGCCCNANGNG 27
||: |::: |:|||||| | | |
Db 7 GATTCTTTTTGTAAGCCCTAGGCG 30

RESULT 15
AAA71126
ID AAA71126 standard; RNA; 42 BP.
XX
AC AAA71126;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #195.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 126; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds

XX
SQ Sequence 42 BP; 12 A; 7 C; 6 G; 0 T; 17 U; 0 Other;

Query Match 62.1%; Score 18; DB 3; Length 42;
Best Local Similarity 75.0%; Pred. No. 2.2;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||| ||||  |||||||| | | |
Db          7 GAUUCUUUUUGUAAGCCCUACGGG 30
```

Search completed: March 23, 2004, 14:53:14
Job time : 236.333 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 14:41:54 ; Search time 55.3333 Seconds
(without alignments)
290.848 Million cell updates/sec

Title: US-09-310-844C-23
Perfect score: 29
Sequence: 1 nnngauncuuunnguaagcccnangngnn 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 895828

Minimum DB seq length: 0
Maximum DB seq length: 70

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	12.8	44.1	68	4	US-08-956-171E-2762	Sequence 2762, Ap
	2	12.2	42.1	27	6	5258283-10	Patent No. 5258283
	3	12.2	42.1	69	2	US-08-410-654B-30	Sequence 30, Appl
	4	12.2	42.1	69	2	US-08-474-851-30	Sequence 30, Appl
	5	12.2	42.1	69	2	US-08-481-560-30	Sequence 30, Appl
c	6	12.2	42.1	70	4	US-08-585-593A-13	Sequence 13, Appl
	7	11.8	40.7	21	2	US-08-747-536-10	Sequence 10, Appl
c	8	11.6	40.0	36	4	US-08-218-369-7	Sequence 7, Appli
	9	11.6	40.0	36	4	US-08-218-369-15	Sequence 15, Appl
c	10	11.6	40.0	36	4	US-09-904-599A-7	Sequence 7, Appli
c	11	11.6	40.0	36	5	PCT-US95-03742-7	Sequence 7, Appli
	12	11.6	40.0	36	5	PCT-US95-03742-15	Sequence 15, Appl
c	13	11.6	40.0	61	4	US-09-619-213B-45	Sequence 45, Appl
c	14	11.4	39.3	65	4	US-09-849-069-37	Sequence 37, Appl
c	15	11.2	38.6	25	1	US-08-741-881-28	Sequence 28, Appl
c	16	11.2	38.6	25	1	US-08-739-158-28	Sequence 28, Appl
c	17	11.2	38.6	25	2	US-08-739-167-28	Sequence 28, Appl
c	18	11.2	38.6	25	3	US-08-404-796-28	Sequence 28, Appl
c	19	11.2	38.6	25	3	US-08-931-869-28	Sequence 28, Appl
c	20	11.2	38.6	25	4	US-09-350-399-28	Sequence 28, Appl
c	21	11.2	38.6	25	4	US-09-236-140A-28	Sequence 28, Appl
	22	11.2	38.6	33	1	US-08-741-881-29	Sequence 29, Appl
	23	11.2	38.6	33	1	US-08-739-158-29	Sequence 29, Appl
	24	11.2	38.6	33	2	US-08-739-167-29	Sequence 29, Appl
	25	11.2	38.6	33	3	US-08-404-796-29	Sequence 29, Appl
	26	11.2	38.6	33	3	US-08-931-869-29	Sequence 29, Appl
	27	11.2	38.6	33	4	US-09-350-399-29	Sequence 29, Appl
	28	11.2	38.6	33	4	US-09-236-140A-29	Sequence 29, Appl
c	29	11.2	38.6	36	2	US-08-642-045B-17	Sequence 17, Appl
c	30	11.2	38.6	36	3	US-08-852-268-17	Sequence 17, Appl
	31	11.2	38.6	41	4	US-09-571-774-2	Sequence 2, Appli
	32	11.2	38.6	41	4	US-09-852-385-2	Sequence 2, Appli
	33	11.2	38.6	47	4	US-09-422-978-2842	Sequence 2842, Ap
	34	11.2	38.6	52	4	US-09-310-463-6	Sequence 6, Appli
	35	11.2	38.6	52	4	US-08-842-248A-6	Sequence 6, Appli
	36	11.2	38.6	70	3	US-09-364-380-29	Sequence 29, Appl
c	37	11	37.9	31	1	US-08-323-531-71	Sequence 71, Appl
c	38	11	37.9	31	1	US-08-198-094-71	Sequence 71, Appl
c	39	11	37.9	31	3	US-08-480-640A-119	Sequence 119, App
c	40	11	37.9	31	3	US-08-295-802-119	Sequence 119, App
c	41	11	37.9	31	3	US-08-107-794A-71	Sequence 71, Appl
c	42	11	37.9	31	3	US-08-488-237A-119	Sequence 119, App
c	43	11	37.9	31	4	US-08-375-992A-119	Sequence 119, App
c	44	11	37.9	31	4	US-08-472-679H-119	Sequence 119, App
c	45	11	37.9	31	5	PCT-US93-07424-71	Sequence 71, Appl

ALIGNMENTS

```
RESULT 1
US-08-956-171E-2762
; Sequence 2762, Application US/08956171E
; Patent No. 6593114
;   GENERAL INFORMATION:
;        APPLICANT: Charles Kunsch
;                   Gil H. Choi
;                   Patrick S. Dillon
;                   Craig A. Rosen
;                   Steven C. Barash
;                   Michael R. Fannon
;        TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
;        NUMBER OF SEQUENCES: 5256
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;             COMPUTER: HP Vectra 486/33
;             OPERATING SYSTEM: MSDOS version 6.2
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/956,171E
;             FILING DATE: 20-Oct-1997
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 60/009,861
;             FILING DATE: January 5, 1996
;             APPLICATION NUMBER: 08/781,986
;             FILING DATE: January 3, 1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Mark J. Hyman
;             REGISTRATION NUMBER: 46,789
;             REFERENCE/DOCKET NUMBER: PB248P1
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (240) 314-1224
;             TELEFAX: (301) 309-8439
;   INFORMATION FOR SEQ ID NO: 2762:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 68 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 2762:
US-08-956-171E-2762
```

```
  Query Match             44.1%;  Score 12.8;  DB 4;  Length 68;
  Best Local Similarity   47.6%;  Pred. No. 1.7e+02;
  Matches   10;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          5 AUNCUUUNNGUAAGCCCNANG 25
              |: |: :   :||||||  | |
Db          3 ATCCTGTTCCTAAGCCCGACG 23


RESULT 2
5258283-10
;Patent No. 5258283
;    APPLICANT: FRAZIER, MARVIN E.;MALLAVIA, LOUIS P.;SAMUEL,
;JAMES E.;BACA, OSWALD G.
;    TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF COXIELLA
;BURNETII IN BIOLOGICAL FLUIDS
;    NUMBER OF SEQUENCES: 17
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/425,856
;      FILING DATE: 23-OCT-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 927,779
;      FILING DATE: 05-NOV-1986
;      APPLICATION NUMBER: 795,207
;      FILING DATE: 05-NOV-1985
;SEQ ID NO:10:
;      LENGTH: 27
5258283-10

  Query Match             42.1%;  Score 12.2;  DB 6;  Length 27;
  Best Local Similarity   40.9%;  Pred. No. 3.1e+02;
  Matches    9;  Conservative    5;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANG 25
              | : |:::   :|||||  | |
Db          4 GGTTCTTTGAATAAGCCAAATG 25


RESULT 3
US-08-410-654B-30
; Sequence 30, Application US/08410654B
; Patent No. 5833976
;  GENERAL INFORMATION:
;    APPLICANT:  Rene de Waal Malefyt
;    APPLICANT:  Di-Hwei Hsu
;    APPLICANT:  Anne O'Garra
;    APPLICANT:  Hergen Spits
;    TITLE OF INVENTION:  Use of Interleukin-10 to Treat
;    TITLE OF INVENTION:  Septic Shock
;    NUMBER OF SEQUENCES:  61
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Schering-Plough Corporation
;      STREET:  2000 Galloping Hill Road
;      CITY:  Kenilworth
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07033
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Macintosh
;      OPERATING SYSTEM:  7.5.3
;      SOFTWARE:  Microsoft Word 5.1a
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/410,654B
;      FILING DATE:  24-MAR-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/229,854
;      FILING DATE:  19-APR-1994
;      APPLICATION NUMBER:  US 07/926,853
;      FILING DATE:  06-AUG-1992
;      APPLICATION NUMBER:  US 07/742,129
;      FILING DATE:  06-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Foulke, Cynthia L.
;      REGISTRATION NUMBER:  32,364
;      REFERENCE/DOCKET NUMBER:  DX0221KQ1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-298-2987
;      TELEFAX:  908-298-5388
;  INFORMATION FOR SEQ ID NO:  30:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  69 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (oligonucleotide)
US-08-410-654B-30

  Query Match             42.1%;  Score 12.2;  DB 2;  Length 69;
  Best Local Similarity   43.5%;  Pred. No. 3.8e+02;
  Matches   10;  Conservative    4;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          5 AUNCUUUNNGUAAGCCCNANGNG 27
              |: | ::   :|||| | | | |
Db         11 ATGCCTTTAATAAGCTCCAAGAG 33


RESULT 4
US-08-474-851-30
; Sequence 30, Application US/08474851
; Patent No. 5837232
;  GENERAL INFORMATION:
;    APPLICANT:  Rene de Waal Malefyt
;    APPLICANT:  Di-Hwei Hsu
;    APPLICANT:  Anne O'Garra
;    APPLICANT:  Hergen Spits
;    TITLE OF INVENTION:  Use of An Interleukin-10 Antagonist to Treat
;    TITLE OF INVENTION:  A B Cell Mediated Autoimmune Disorder
;    NUMBER OF SEQUENCES:  61
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Schering-Plough Corporation
;      STREET:  2000 Galloping Hill Road
;      CITY:  Kenilworth
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07033
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Macintosh
;      OPERATING SYSTEM:  7.5.3
;      SOFTWARE:  Microsoft Word 6.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/474,851
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/410,654
;      FILING DATE:  24-MAR-1995
;      APPLICATION NUMBER:  US 08/229,854
;      FILING DATE:  19-APR-1994
;      APPLICATION NUMBER:  US 07/926,853
;      FILING DATE:  06-AUG-1992
;      APPLICATION NUMBER:  US 07/742,129
;      FILING DATE:  06-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Foulke, Cynthia L.
;      REGISTRATION NUMBER:  32,364
;      REFERENCE/DOCKET NUMBER:  DX0221KQ1GD
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-298-2987
;      TELEFAX:  908-298-5388
;  INFORMATION FOR SEQ ID NO:  30:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  69 base pairs
;      TYPE:  nucleic acid
```

```
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (oligonucleotide)
US-08-474-851-30

  Query Match             42.1%;  Score 12.2;  DB 2;  Length 69;
  Best Local Similarity   43.5%;  Pred. No. 3.8e+02;
  Matches   10;  Conservative    4;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          5 AUNCUUUNNGUAAGCCCNANGNG 27
              |: | ::   :|||| | | | |
Db         11 ATGCCTTTAATAAGCTCCAAGAG 33


RESULT 5
US-08-481-560-30
; Sequence 30, Application US/08481560
; Patent No. 5837293
;  GENERAL INFORMATION:
;    APPLICANT:  Rene de Waal Malefyt
;    APPLICANT:  Di-Hwei Hsu
;    APPLICANT:  Anne O'Garra
;    APPLICANT:  Hergen Spits
;    TITLE OF INVENTION:  Use of Interleukin-10 to Modulate
;    TITLE OF INVENTION:  Inflammation or T-Cell Mediated
;    TITLE OF INVENTION:  Immune Function
;    NUMBER OF SEQUENCES:  61
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Schering-Plough Corporation
;      STREET:  2000 Galloping Hill Road
;      CITY:  Kenilworth
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07033
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Macintosh
;      OPERATING SYSTEM:  7.5.3
;      SOFTWARE:  Microsoft Word 6.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/481,560
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/410,654
;      FILING DATE:  24-MAR-1995
;      APPLICATION NUMBER:  US 08/229,854
;      FILING DATE:  19-APR-1994
;      APPLICATION NUMBER:  US 07/926,853
;      FILING DATE:  06-AUG-1992
;      APPLICATION NUMBER:  US 07/742,129
;      FILING DATE:  06-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Foulke, Cynthia L.
;      REGISTRATION NUMBER:  32,364
;      REFERENCE/DOCKET NUMBER:  DX0221KQ1GC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-298-2987
;      TELEFAX:  908-298-5388
;  INFORMATION FOR SEQ ID NO:  30:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  69 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (oligonucleotide)
US-08-481-560-30

  Query Match             42.1%;  Score 12.2;  DB 2;  Length 69;
  Best Local Similarity   43.5%;  Pred. No. 3.8e+02;
  Matches   10;  Conservative    4;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          5 AUNCUUUNNGUAAGCCCNANGNG 27
              |: | ::   :|||| | | | |
Db         11 ATGCCTTTAATAAGCTCCAAGAG 33


RESULT 6
US-08-585-593A-13/c
; Sequence 13, Application US/08585593A
; Patent No. 6503706
;  GENERAL INFORMATION:
;    APPLICANT:  ABKEN, Hinrich J
;    APPLICANT:  ALBERT, Winfried
;    APPLICANT:  JUNGFER, Herbert
;    TITLE OF INVENTION:  METHOD OF IDENTIFYING HUMAN AND ANIMAL
;    TITLE OF INVENTION:  CELLS CAPABLE OF UNLIMITED PROLIFERATION OR TUMOR
;    TITLE OF INVENTION:  FORMATION
;    NUMBER OF SEQUENCES:  66
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Nikaido, Marmelstein, Murray & Oram LLP
;      STREET:  655 Fifteenth Street N.W. Suite 330
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/585,593A
;      FILING DATE:  16-JAN-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP94/02307
;      FILING DATE:  13-JUL-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE P 43 23 727.4
;      FILING DATE:  15-JUL-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murray, Robert B.
;      REGISTRATION NUMBER:  22,980
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)638-5000
;      TELEFAX:  (202)638-4810
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  70 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-08-585-593A-13

  Query Match             42.1%;  Score 12.2;  DB 4;  Length 70;
  Best Local Similarity   40.9%;  Pred. No. 3.9e+02;
  Matches    9;  Conservative    5;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANG 25
              ||: |:::  |:|  ||  | |
Db         70 GATCCTTTCGGTATTCCAGAAG 49


RESULT 7
US-08-747-536-10
; Sequence 10, Application US/08747536
; Patent No. 5968737
;  GENERAL INFORMATION:
;    APPLICANT:  Ali-Osman, Francis
;    APPLICANT:  Lopez-Berestein, Gabriel
;    APPLICANT:  Buolamwini, John
;    APPLICANT:  Antoun, Gamil
;    APPLICANT:  Lo, Hui-Wen
;    APPLICANT:  Keller, Charles
```

```
;    APPLICANT:  Akande, Olanike
;    TITLE OF INVENTION:  GLUTATHIONE S-TRANSFERASE (GST) GENES IN
;    TITLE OF INVENTION:  CANCER
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/747,536
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  UTXC:492
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/418-3000
;      TELEFAX:  512/474-7577
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-747-536-10

  Query Match             40.7%;  Score 11.8;  DB 2;  Length 21;
  Best Local Similarity   50.0%;  Pred. No. 5e+02;
  Matches    9;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCC 21
              ||  |:::  |: |||||
Db          2 GAGGCTTTGAGTGAGCCC 19


RESULT 8
US-08-218-369-7/c
; Sequence 7, Application US/08218369
; Patent No. 6312699
;  GENERAL INFORMATION:
;    APPLICANT:  Curiel, David T.
;    APPLICANT:  Engler, Jeffrey A.
;    TITLE OF INVENTION:  Ligands Added to Adenovirus Fiber
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patrea L. Pabst
;      STREET:  1100 Peachtree Street, Suite 2800
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30309-4530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/218,369
;      FILING DATE:  28-MAR-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  IGI101
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404) 815-6508
;      TELEFAX:  (404) 815-6555
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  36 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..36
;      OTHER INFORMATION:  /note= "Nucleotide sequence encoding a streptavidin mimic
US-08-218-369-7

  Query Match             40.0%;  Score 11.6;  DB 4;  Length 36;
  Best Local Similarity   41.7%;  Pred. No. 7.6e+02;
  Matches   10;  Conservative    4;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||  |:::  |:  | || | | |
Db         31 GAAGCTTTAGGTGGGGCCCATGAG 8


RESULT 9
US-08-218-369-15
; Sequence 15, Application US/08218369
; Patent No. 6312699
;  GENERAL INFORMATION:
;    APPLICANT:  Curiel, David T.
;    APPLICANT:  Engler, Jeffrey A.
;    TITLE OF INVENTION:  Ligands Added to Adenovirus Fiber
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patrea L. Pabst
;      STREET:  1100 Peachtree Street, Suite 2800
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30309-4530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/218,369
;      FILING DATE:  28-MAR-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  IGI101
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404) 815-6508
;      TELEFAX:  (404) 815-6555
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  36 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..36
;      OTHER INFORMATION:  /note= "Nucleotides 5 through 36 are complementary to nucl
US-08-218-369-15
```

```
  Query Match             40.0%;  Score 11.6;  DB 4;  Length 36;
  Best Local Similarity   41.7%;  Pred. No. 7.6e+02;
  Matches   10;  Conservative    4;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||  |:::  |:  | || | | |
Db         10 GAAGCTTTAGGTGGGGCCCATGAG 33


RESULT 10
US-09-904-599A-7/c
; Sequence 7, Application US/09904599A
; Patent No. 6683170
; GENERAL INFORMATION:
;  APPLICANT: Curiel, David T.,
;  APPLICANT:  Engler, Jeffrey A.
;  TITLE OF INVENTION: Ligands Added to Adenovirus Fiber
;  FILE REFERENCE: D5839/D
;  CURRENT APPLICATION NUMBER: US/09/904,599A
;  CURRENT FILING DATE:  2001-07-13
;  PRIOR APPLICATION NUMBER: US 08/218,369
;  PRIOR FILING DATE: 1994-03-28
;  NUMBER OF SEQ ID NOS: 13
; SEQ ID NO 7
;   LENGTH: 36
;   TYPE: DNA
;   ORGANISM: artificial sequence
;   FEATURE:
;   OTHER INFORMATION: sequence coding for streptavidin mimic
;   OTHER INFORMATION: that binds biotin nucleotide sequence
US-09-904-599A-7

  Query Match             40.0%;  Score 11.6;  DB 4;  Length 36;
  Best Local Similarity   41.7%;  Pred. No. 7.6e+02;
  Matches   10;  Conservative    4;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||  |:::  |:  | || | | |
Db         31 GAAGCTTTAGGTGGGGCCCATGAG 8


RESULT 11
PCT-US95-03742-7/c
; Sequence 7, Application PC/TUS9503742
;  GENERAL INFORMATION:
;    APPLICANT:  The UAB Research Foundation
;    TITLE OF INVENTION:  Ligands Added to Adenovirus Fiber
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patrea L. Pabst
;      STREET:  2800 One Atlantic Center
;      STREET:  1201 West Peachtree Street
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30309-3450
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/03742
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  IGI101
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404) 873-8794
;      TELEFAX:  (404) 873-8795
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  36 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..36
;      OTHER INFORMATION:  /note= "Nucleotide sequence
;      OTHER INFORMATION:  encoding a streptavidin mimic that binds biotin."
PCT-US95-03742-7

  Query Match             40.0%;  Score 11.6;  DB 5;  Length 36;
  Best Local Similarity   41.7%;  Pred. No. 7.6e+02;
  Matches   10;  Conservative    4;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||  |:::  |:  | || | | |
Db         31 GAAGCTTTAGGTGGGGCCCATGAG 8


RESULT 12
PCT-US95-03742-15
; Sequence 15, Application PC/TUS9503742
;  GENERAL INFORMATION:
;    APPLICANT:  The UAB Research Foundation
;    TITLE OF INVENTION:  Ligands Added to Adenovirus Fiber
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patrea L. Pabst
;      STREET:  2800 One Atlantic Center
;      STREET:  1201 West Peachtree Street
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30309-3450
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/03742
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  IGI101
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404) 873-8794
;      TELEFAX:  (404) 873-8795
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  36 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..36
;      OTHER INFORMATION:  /note= "Nucleotides 5 through 36
;      OTHER INFORMATION:  are complementary to nucleotides 5 through 36 of
;      OTHER INFORMATION:  Sequence ID No. 7."
PCT-US95-03742-15
```

```
  Query Match             40.0%;  Score 11.6;  DB 5;  Length 36;
  Best Local Similarity   41.7%;  Pred. No. 7.6e+02;
  Matches   10;  Conservative    4;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||  |:::  |:  | || | | |
Db         10 GAAGCTTTAGGTGGGGCCCATGAG 33


RESULT 13
US-09-619-213B-45/c
; Sequence 45, Application US/09619213B
; Patent No. 6458539
; GENERAL INFORMATION:
;  APPLICANT: Gold, Larry
;  APPLICANT:  Smith, Jonathan Drew
;  APPLICANT:  Koch, Tad
;  APPLICANT:  Golden, Mace
;  TITLE OF INVENTION: Photoselection of Nucleic Acid Ligands
;  FILE REFERENCE: NEX10-5
;  CURRENT APPLICATION NUMBER: US/09/619,213B
;  CURRENT FILING DATE:  2000-07-19
;  PRIOR APPLICATION NUMBER: 09/459,553
;  PRIOR FILING DATE: 1999-12-13
;  PRIOR APPLICATION NUMBER: 09/093,293
;  PRIOR FILING DATE: 1998-06-08
;  PRIOR APPLICATION NUMBER: 08/612,895
;  PRIOR FILING DATE: 1996-03-08
;  PRIOR APPLICATION NUMBER: 08/123,935
;  PRIOR FILING DATE: 1993-09-17
;  NUMBER OF SEQ ID NOS: 100
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
;   LENGTH: 61
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Nucleic Acid
;   OTHER INFORMATION: Ligand
;   NAME/KEY: modified_base
;   LOCATION: (1)..(61)
;   OTHER INFORMATION: All T's are 5-bromouracil
US-09-619-213B-45

  Query Match             40.0%;  Score 11.6;  DB 4;  Length 61;
  Best Local Similarity   45.8%;  Pred. No. 8.6e+02;
  Matches   11;  Conservative    3;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANGNG 27
              ||:.|: :    ||||||  |    |
Db         42 GATACTATGAACAAGCCCCAGTGG 19


RESULT 14
US-09-849-069-37/c
; Sequence 37, Application US/09849069
; Patent No. 6630306
; GENERAL INFORMATION:
;  APPLICANT: Ronald R. Breaker
;  TITLE OF INVENTION: Bioreactive Allosteric Polynucleotides
;  FILE REFERENCE: OCR-794.CIP
;  CURRENT APPLICATION NUMBER: US/09/849,069
;  CURRENT FILING DATE:  2001-05-07
;  PRIOR APPLICATION NUMBER: US 09/331,809
;  PRIOR FILING DATE: 1999-06-18
;  NUMBER OF SEQ ID NOS: 57
;  SOFTWARE: MS-DOS
; SEQ ID NO 37
;   LENGTH: 65
;   TYPE: DNA
;   ORGANISM: artificial sequence
;   FEATURE:
;   OTHER INFORMATION: DNA with 3 cleavage sites
US-09-849-069-37

  Query Match             39.3%;  Score 11.4;  DB 4;  Length 65;
  Best Local Similarity   50.0%;  Pred. No. 1.2e+03;
  Matches    9;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy         10 UUNNGUAAGCCCNANGNG 27
              ::  |:|||||  | | |
Db         51 TTCGGTAAGCCACATGAG 34


RESULT 15
US-08-741-881-28/c
; Sequence 28, Application US/08741881
; Patent No. 5789245
;  GENERAL INFORMATION:
;    APPLICANT:  Dubensky Jr, Thomas W
;    APPLICANT:  Polo, John M.
;    APPLICANT:  Ibanez, Carlos E.
;    APPLICANT:  Chang, Stephen M.W.
;    APPLICANT:  Jolly, Douglas J.
;    APPLICANT:  Driver, David A.
;    APPLICANT:  Belli, Barbara A.
;    TITLE OF INVENTION:  EUKARYOTIC LAYERED VECTOR INITIATION SYSTEMS
;    NUMBER OF SEQUENCES:  128
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/741,881
;      FILING DATE:  30-OCT-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  930049.423C6 / 1146.007
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  25 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-741-881-28

  Query Match             38.6%;  Score 11.2;  DB 1;  Length 25;
  Best Local Similarity   36.4%;  Pred. No. 1.2e+03;
  Matches    8;  Conservative    5;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          6 UNCUUUNNGUAAGCCCNANGNG 27
              : |:::  |: ||||  |    |
Db         24 TCCTTTAGGTTAGCCGTACAAG 3


Search completed: March 23, 2004, 17:20:31
Job time : 60.3333 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 14:06:04 ; Search time 237.667 Seconds
(without alignments)
451.369 Million cell updates/sec

Title: US-09-310-844C-23
Perfect score: 29
Sequence: 1 nnngauncuuunnguaagcccnangngnn 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2438257 seqs, 1849576744 residues

Total number of hits satisfying chosen parameters: 1514106

Minimum DB seq length: 0
Maximum DB seq length: 70

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	12.8	44.1	68	8	US-08-781-986A-2762	Sequence 2762, Ap
	2	12.8	44.1	68	12	US-10-329-624-2762	Sequence 2762, Ap
c	3	12.4	42.8	60	10	US-09-908-975-18725	Sequence 18725, A
c	4	12.4	42.8	65	10	US-09-908-975-2848	Sequence 2848, Ap
	5	12.2	42.1	50	15	US-10-131-827-464	Sequence 464, App
	6	12.2	42.1	60	10	US-09-908-975-9828	Sequence 9828, Ap
c	7	12.2	42.1	60	10	US-09-908-975-12187	Sequence 12187, A
c	8	12.2	42.1	60	10	US-09-908-975-15109	Sequence 15109, A
	9	12.2	42.1	60	10	US-09-908-975-18934	Sequence 18934, A
	10	11.8	40.7	25	14	US-10-098-263B-76444	Sequence 76444, A
	11	11.8	40.7	47	9	US-09-230-926A-35	Sequence 35, Appl
c	12	11.8	40.7	60	10	US-09-908-975-15914	Sequence 15914, A
	13	11.8	40.7	60	10	US-09-908-975-17626	Sequence 17626, A
	14	11.8	40.7	65	10	US-09-908-975-1254	Sequence 1254, Ap
	15	11.8	40.7	65	10	US-09-908-975-30297	Sequence 30297, A
	16	11.8	40.7	65	14	US-10-032-585-316	Sequence 316, App
	17	11.6	40.0	24	10	US-09-964-895-27	Sequence 27, Appl
	18	11.6	40.0	24	14	US-10-059-152-26	Sequence 26, Appl
c	19	11.6	40.0	31	10	US-09-848-754A-6937	Sequence 6937, Ap
c	20	11.6	40.0	31	10	US-09-848-754A-7188	Sequence 7188, Ap
c	21	11.6	40.0	31	10	US-09-848-754A-7495	Sequence 7495, Ap
c	22	11.6	40.0	31	10	US-09-740-332-6639	Sequence 6639, Ap
c	23	11.6	40.0	31	10	US-09-740-332-9154	Sequence 9154, Ap
c	24	11.6	40.0	31	10	US-09-817-879-6639	Sequence 6639, Ap
c	25	11.6	40.0	31	10	US-09-817-879-9154	Sequence 9154, Ap
c	26	11.6	40.0	31	14	US-10-163-552-1019	Sequence 1019, Ap
c	27	11.6	40.0	31	14	US-10-156-306-3281	Sequence 3281, Ap
c	28	11.6	40.0	36	9	US-09-904-599A-7	Sequence 7, Appli
	29	11.6	40.0	38	15	US-10-388-329-15	Sequence 15, Appl
c	30	11.6	40.0	56	10	US-09-800-130A-8	Sequence 8, Appli
c	31	11.6	40.0	56	14	US-10-413-909-8	Sequence 8, Appli
	32	11.6	40.0	60	10	US-09-908-975-5781	Sequence 5781, Ap
c	33	11.6	40.0	60	10	US-09-908-975-12753	Sequence 12753, A
c	34	11.6	40.0	60	10	US-09-908-975-14781	Sequence 14781, A
	35	11.6	40.0	65	10	US-09-908-975-24835	Sequence 24835, A
	36	11.6	40.0	65	10	US-09-908-975-29918	Sequence 29918, A
	37	11.4	39.3	21	15	US-10-435-696-279	Sequence 279, App
c	38	11.4	39.3	25	14	US-10-098-263B-5191	Sequence 5191, Ap
c	39	11.4	39.3	25	14	US-10-098-263B-5192	Sequence 5192, Ap
	40	11.4	39.3	25	14	US-10-098-263B-94421	Sequence 94421, A
	41	11.4	39.3	44	14	US-10-207-655-28	Sequence 28, Appl
	42	11.4	39.3	44	14	US-10-053-530-28	Sequence 28, Appl
	43	11.4	39.3	60	10	US-09-908-975-6202	Sequence 6202, Ap
c	44	11.4	39.3	60	10	US-09-908-975-7920	Sequence 7920, Ap
c	45	11.4	39.3	60	10	US-09-908-975-20533	Sequence 20533, A

ALIGNMENTS

RESULT 1
US-08-781-986A-2762
; Sequence 2762, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 2762:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 68 base pairs
; TYPE: nucleic acid

```
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
US-08-781-986A-2762

  Query Match             44.1%;  Score 12.8;  DB 8;  Length 68;
  Best Local Similarity   47.6%;  Pred. No. 2.1e+03;
  Matches   10;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          5 AUNCUUUNNGUAAGCCCNANG 25
              |: |: :   :|||||| | |
Db          3 ATCCTGTTCCTAAGCCCGACG 23


RESULT 2
US-10-329-624-2762
; Sequence 2762, Application US/10329624
; Publication No. US20040043037A1
;   GENERAL INFORMATION:
;        APPLICANT: Charles Kunsch
;                   Gil H. Choi
;                   Patrick S. Dillon
;                   Craig A. Rosen
;                   Steven C. Barash
;                   Michael R. Fannon
;        TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
;        NUMBER OF SEQUENCES: 5256
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;             COMPUTER: HP Vectra 486/33
;             OPERATING SYSTEM: MSDOS version 6.2
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/329,624
;             FILING DATE: 27-Dec-2002
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/956,171
;             FILING DATE: October 20, 1997
;             APPLICATION NUMBER: 60/009,861
;             FILING DATE: January 5, 1996
;             APPLICATION NUMBER: 08/781,986
;             FILING DATE: January 3, 1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Mark J. Hyman
;             REGISTRATION NUMBER: 46,789
;             REFERENCE/DOCKET NUMBER: PB248P1D1
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (240) 314-1224
;             TELEFAX: (301) 309-8439
;   INFORMATION FOR SEQ ID NO: 2762:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 68 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 2762:
US-10-329-624-2762

  Query Match             44.1%;  Score 12.8;  DB 12;  Length 68;
  Best Local Similarity   47.6%;  Pred. No. 2.1e+03;
  Matches   10;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          5 AUNCUUUNNGUAAGCCCNANG 25
              |: |: :   :|||||| | |
Db          3 ATCCTGTTCCTAAGCCCGACG 23


RESULT 3
US-09-908-975-18725/c
; Sequence 18725, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
;  APPLICANT: SHOSHAN, Avi
;  APPLICANT:  WASSERMAN, Alon
;  APPLICANT:  MINTZ, Eli
;  APPLICANT:  MINTZ, Liat
;  APPLICANT:  FAIGLER, Simchon
;  TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE \
;  TITLE OF INVENTION:  THAT POPULATE A TRANSCRIPTOME
;  FILE REFERENCE: 36688-0005
;  CURRENT APPLICATION NUMBER: US/09/908,975
;  CURRENT FILING DATE:  2001-07-20
;  PRIOR APPLICATION NUMBER: US 60/287,724
;  PRIOR FILING DATE: 2001-05-02
;  PRIOR APPLICATION NUMBER: US 60/221,607
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 32337
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18725
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-908-975-18725

  Query Match             42.8%;  Score 12.4;  DB 10;  Length 60;
  Best Local Similarity   55.6%;  Pred. No. 3.6e+03;
  Matches   10;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          8 CUUUNNGUAAGCCCNANG 25
              |:::  | ||||||  | |
Db         25 CTTTCTGAAAGCCCCATG 8


RESULT 4
US-09-908-975-2848/c
; Sequence 2848, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
;  APPLICANT: SHOSHAN, Avi
;  APPLICANT:  WASSERMAN, Alon
;  APPLICANT:  MINTZ, Eli
;  APPLICANT:  MINTZ, Liat
;  APPLICANT:  FAIGLER, Simchon
;  TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE \
;  TITLE OF INVENTION:  THAT POPULATE A TRANSCRIPTOME
;  FILE REFERENCE: 36688-0005
;  CURRENT APPLICATION NUMBER: US/09/908,975
;  CURRENT FILING DATE:  2001-07-20
;  PRIOR APPLICATION NUMBER: US 60/287,724
;  PRIOR FILING DATE: 2001-05-02
;  PRIOR APPLICATION NUMBER: US 60/221,607
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 32337
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2848
;   LENGTH: 65
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
US-09-908-975-2848

  Query Match             42.8%;  Score 12.4;  DB 10;  Length 65;
  Best Local Similarity   52.9%;  Pred. No. 3.6e+03;
  Matches    9;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCC 20
              ||: |::   |:|||||
Db         53 GATACTTGCAGTAAGCC 37


RESULT 5
```

```
US-10-131-827-464
; Sequence 464, Application US/10131827
; Publication No. US20040009479A1
; GENERAL INFORMATION:
;  APPLICANT: Wohlgemuth, Jay
;  APPLICANT:  Fry, Kirk
;  APPLICANT:  Woodward, Robert
;  APPLICANT:  Ly, Ngoc
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE
;  TITLE OF INVENTION:  CHRONIC INFLAMMATORY DISEASES
;  FILE REFERENCE: 506612000120
;  CURRENT APPLICATION NUMBER: US/10/131,827
;  CURRENT FILING DATE:  2002-09-06
;  PRIOR APPLICATION NUMBER: US 10/006,290
;  PRIOR FILING DATE: 2001-10-22
;  PRIOR APPLICATION NUMBER: US 60/296,764
;  PRIOR FILING DATE: 2001-06-08
;  NUMBER OF SEQ ID NOS: 9090
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 464
;   LENGTH: 50
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-131-827-464

  Query Match             42.1%;  Score 12.2;  DB 15;  Length 50;
  Best Local Similarity   45.5%;  Pred. No. 4.5e+03;
  Matches   10;  Conservative    4;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANG 25
              ||  |:::   :||| || | |
Db         22 GAGGCTTTTCCTAAGGCCAAGG 43


RESULT 6
US-09-908-975-9828
; Sequence 9828, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
;  APPLICANT: SHOSHAN, Avi
;  APPLICANT:  WASSERMAN, Alon
;  APPLICANT:  MINTZ, Eli
;  APPLICANT:  MINTZ, Liat
;  APPLICANT:  FAIGLER, Simchon
;  TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
;  TITLE OF INVENTION:  THAT POPULATE A TRANSCRIPTOME
;  FILE REFERENCE: 36688-0005
;  CURRENT APPLICATION NUMBER: US/09/908,975
;  CURRENT FILING DATE:  2001-07-20
;  PRIOR APPLICATION NUMBER: US 60/287,724
;  PRIOR FILING DATE: 2001-05-02
;  PRIOR APPLICATION NUMBER: US 60/221,607
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 32337
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9828
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-908-975-9828

  Query Match             42.1%;  Score 12.2;  DB 10;  Length 60;
  Best Local Similarity   52.2%;  Pred. No. 4.7e+03;
  Matches   12;  Conservative    2;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          5 AUNCUUUNNGUAAGCCCNANGNG 27
              |  |:    |:|||||| | | |
Db         34 AACCTCATGGTAAGCCCAACGTG 56


RESULT 7
US-09-908-975-12187/c
; Sequence 12187, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
;  APPLICANT: SHOSHAN, Avi
;  APPLICANT:  WASSERMAN, Alon
;  APPLICANT:  MINTZ, Eli
;  APPLICANT:  MINTZ, Liat
;  APPLICANT:  FAIGLER, Simchon
;  TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
;  TITLE OF INVENTION:  THAT POPULATE A TRANSCRIPTOME
;  FILE REFERENCE: 36688-0005
;  CURRENT APPLICATION NUMBER: US/09/908,975
;  CURRENT FILING DATE:  2001-07-20
;  PRIOR APPLICATION NUMBER: US 60/287,724
;  PRIOR FILING DATE: 2001-05-02
;  PRIOR APPLICATION NUMBER: US 60/221,607
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 32337
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12187
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-908-975-12187

  Query Match             42.1%;  Score 12.2;  DB 10;  Length 60;
  Best Local Similarity   40.9%;  Pred. No. 4.7e+03;
  Matches    9;  Conservative    5;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANG 25
              ||: |:::  |:|  | | | |
Db         54 GATTCTTTCTGTAGCCGCTAAG 33


RESULT 8
US-09-908-975-15109/c
; Sequence 15109, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
;  APPLICANT: SHOSHAN, Avi
;  APPLICANT:  WASSERMAN, Alon
;  APPLICANT:  MINTZ, Eli
;  APPLICANT:  MINTZ, Liat
;  APPLICANT:  FAIGLER, Simchon
;  TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
;  TITLE OF INVENTION:  THAT POPULATE A TRANSCRIPTOME
;  FILE REFERENCE: 36688-0005
;  CURRENT APPLICATION NUMBER: US/09/908,975
;  CURRENT FILING DATE:  2001-07-20
;  PRIOR APPLICATION NUMBER: US 60/287,724
;  PRIOR FILING DATE: 2001-05-02
;  PRIOR APPLICATION NUMBER: US 60/221,607
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 32337
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15109
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-908-975-15109

  Query Match             42.1%;  Score 12.2;  DB 10;  Length 60;
  Best Local Similarity   45.5%;  Pred. No. 4.7e+03;
  Matches   10;  Conservative    4;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANG 25
              ||: |:::  | | | || | |
Db         37 GATGCTTTCTGCATGGCCAAGG 16


RESULT 9
US-09-908-975-18934
; Sequence 18934, Application US/09908975
; Publication No. US20030165843A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: SHOSHAN, Avi
;  APPLICANT:  WASSERMAN, Alon
;  APPLICANT:  MINTZ, Eli
;  APPLICANT:  MINTZ, Liat
;  APPLICANT:  FAIGLER, Simchon
;  TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
;  TITLE OF INVENTION:  THAT POPULATE A TRANSCRIPTOME
;  FILE REFERENCE: 36688-0005
;  CURRENT APPLICATION NUMBER: US/09/908,975
;  CURRENT FILING DATE:  2001-07-20
;  PRIOR APPLICATION NUMBER: US 60/287,724
;  PRIOR FILING DATE: 2001-05-02
;  PRIOR APPLICATION NUMBER: US 60/221,607
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 32337
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18934
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-908-975-18934

  Query Match             42.1%;  Score 12.2;  DB 10;  Length 60;
  Best Local Similarity   43.5%;  Pred. No. 4.7e+03;
  Matches   10;  Conservative    4;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          5 AUNCUUUNNGUAAGCCCNANGNG 27
              |: | ::  |:|||| |   | |
Db         27 ATCCGTTGCGTAAGCACAGAGAG 49


RESULT 10
US-10-098-263B-76444
; Sequence 76444, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
;  APPLICANT: Mittman, Michael
;  TITLE OF INVENTION: Human Microarray
;  FILE REFERENCE: 3118.1
;  CURRENT APPLICATION NUMBER: US/10/098,263B
;  CURRENT FILING DATE:  2003-01-08
;  PRIOR APPLICATION NUMBER: 60/276,759
;  PRIOR FILING DATE: 2001-03-16
;  NUMBER OF SEQ ID NOS: 131066
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 76444
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-098-263B-76444

  Query Match             40.7%;  Score 11.8;  DB 14;  Length 25;
  Best Local Similarity   44.4%;  Pred. No. 6.9e+03;
  Matches    8;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCC 21
              ||: |:::   :||| ||
Db          1 GATACTTTTTTTAAGTCC 18


RESULT 11
US-09-230-926A-35
; Sequence 35, Application US/09230926A
; Patent No. US20020168633A1
; GENERAL INFORMATION:
;  APPLICANT: MABILAT, Claude
;  APPLICANT:  SCHLEIFER, Karl-Heinz
;  APPLICANT:  LUDWIG, Wolfgang
;  TITLE OF INVENTION: NUCLEOTIDE FRAGMENT OF THE 23S RNA OF BACTERIA OF THE GENUS CHLAM
;  TITLE OF INVENTION:  USE AS A PROBE, PRIMER, AND IN A REAGENT AND A DETECTION PROCEDU
;  FILE REFERENCE: 102682
;  CURRENT APPLICATION NUMBER: US/09/230,926A
;  CURRENT FILING DATE:  1999-03-04
;  PRIOR APPLICATION NUMBER: PCT/FR98/01157
;  PRIOR FILING DATE: 1998-06-05
;  PRIOR APPLICATION NUMBER: FR 97/07200
;  PRIOR FILING DATE: 1997-06-05
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
;   LENGTH: 47
;   TYPE: RNA
;   ORGANISM: Chlamydia pneumoniae
US-09-230-926A-35

  Query Match             40.7%;  Score 11.8;  DB 9;  Length 47;
  Best Local Similarity   65.0%;  Pred. No. 7.6e+03;
  Matches   13;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          6 UNCUUUNNGUAAGCCCNANG 25
              | ||    |||||||| | |
Db         28 UCCUCGCCGUAAGCCCAAGG 47


RESULT 12
US-09-908-975-15914/c
; Sequence 15914, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
;  APPLICANT: SHOSHAN, Avi
;  APPLICANT:  WASSERMAN, Alon
;  APPLICANT:  MINTZ, Eli
;  APPLICANT:  MINTZ, Liat
;  APPLICANT:  FAIGLER, Simchon
;  TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE \
;  TITLE OF INVENTION:  THAT POPULATE A TRANSCRIPTOME
;  FILE REFERENCE: 36688-0005
;  CURRENT APPLICATION NUMBER: US/09/908,975
;  CURRENT FILING DATE:  2001-07-20
;  PRIOR APPLICATION NUMBER: US 60/287,724
;  PRIOR FILING DATE: 2001-05-02
;  PRIOR APPLICATION NUMBER: US 60/221,607
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 32337
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15914
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-908-975-15914

  Query Match             40.7%;  Score 11.8;  DB 10;  Length 60;
  Best Local Similarity   50.0%;  Pred. No. 8e+03;
  Matches    9;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCC 21
              ||  |::: |: |||||
Db         30 GAGGCTTTGAGTGAGCCC 13


RESULT 13
US-09-908-975-17626
; Sequence 17626, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
;  APPLICANT: SHOSHAN, Avi
;  APPLICANT:  WASSERMAN, Alon
;  APPLICANT:  MINTZ, Eli
;  APPLICANT:  MINTZ, Liat
;  APPLICANT:  FAIGLER, Simchon
;  TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE \
;  TITLE OF INVENTION:  THAT POPULATE A TRANSCRIPTOME
;  FILE REFERENCE: 36688-0005
;  CURRENT APPLICATION NUMBER: US/09/908,975
;  CURRENT FILING DATE:  2001-07-20
```

```
;  PRIOR APPLICATION NUMBER: US 60/287,724
;  PRIOR FILING DATE: 2001-05-02
;  PRIOR APPLICATION NUMBER: US 60/221,607
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 32337
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17626
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-908-975-17626

  Query Match             40.7%;  Score 11.8;  DB 10;  Length 60;
  Best Local Similarity   40.0%;  Pred. No. 8e+03;
  Matches    8;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          6 UNCUUUNNGUAAGCCCNANG 25
              : |:::  |:|||| |   |
Db         28 TGCTTTTGGTAAGCACTTTG 47


RESULT 14
US-09-908-975-1254
; Sequence 1254, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
;  APPLICANT: SHOSHAN, Avi
;  APPLICANT:  WASSERMAN, Alon
;  APPLICANT:  MINTZ, Eli
;  APPLICANT:  MINTZ, Liat
;  APPLICANT:  FAIGLER, Simchon
;  TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
;  TITLE OF INVENTION:  THAT POPULATE A TRANSCRIPTOME
;  FILE REFERENCE: 36688-0005
;  CURRENT APPLICATION NUMBER: US/09/908,975
;  CURRENT FILING DATE:  2001-07-20
;  PRIOR APPLICATION NUMBER: US 60/287,724
;  PRIOR FILING DATE: 2001-05-02
;  PRIOR APPLICATION NUMBER: US 60/221,607
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 32337
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1254
;   LENGTH: 65
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
US-09-908-975-1254

  Query Match             40.7%;  Score 11.8;  DB 10;  Length 65;
  Best Local Similarity   45.0%;  Pred. No. 8.1e+03;
  Matches    9;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          6 UNCUUUNNGUAAGCCCNANG 25
              : |: :  |:|||| | | |
Db         11 TGCTCTTGGTAAGCTCCAGG 30


RESULT 15
US-09-908-975-30297
; Sequence 30297, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
;  APPLICANT: SHOSHAN, Avi
;  APPLICANT:  WASSERMAN, Alon
;  APPLICANT:  MINTZ, Eli
;  APPLICANT:  MINTZ, Liat
;  APPLICANT:  FAIGLER, Simchon
;  TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
;  TITLE OF INVENTION:  THAT POPULATE A TRANSCRIPTOME
;  FILE REFERENCE: 36688-0005
;  CURRENT APPLICATION NUMBER: US/09/908,975
;  CURRENT FILING DATE:  2001-07-20
;  PRIOR APPLICATION NUMBER: US 60/287,724
;  PRIOR FILING DATE: 2001-05-02
;  PRIOR APPLICATION NUMBER: US 60/221,607
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 32337
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30297
;   LENGTH: 65
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-908-975-30297

  Query Match             40.7%;  Score 11.8;  DB 10;  Length 65;
  Best Local Similarity   50.0%;  Pred. No. 8.1e+03;
  Matches    9;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCC 21
              ||: |:::    ||||||
Db         37 GATTCTTTCCCCAAGCCC 54


Search completed: March 23, 2004, 17:17:30
Job time : 244 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 14:02:34 ; Search time 1997.33 Seconds
(without alignments)
433.580 Million cell updates/sec

Title: US-09-310-844C-23
Perfect score: 29
Sequence: 1 nnngauncuuunnguaagcccnangngnn 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 289680

Minimum DB seq length: 0
Maximum DB seq length: 70

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	13	44.8	46	28	AZ833686	AZ833686 2M0115L20
	2	12.8	44.1	70	28	BH759592	BH759592 KG05236-3
c	3	12.4	42.8	52	9	AA700959	AA700959 zf87d10.s
c	4	12.4	42.8	70	9	AA468615	AA468615 ne08b04.s
	5	12.2	42.1	48	28	AZ503560	AZ503560 1M0343E21
	6	12.2	42.1	56	28	BZ770420	BZ770420 SALK_1433
	7	12.2	42.1	66	13	BX744082	BX744082 BX744082
c	8	12.2	42.1	67	14	CD946435	CD946435 REN_47 Ge
c	9	11.8	40.7	40	9	AA975071	AA975071 cn03d07.s
c	10	11.8	40.7	49	10	BE970036	BE970036 601680150
	11	11.8	40.7	51	29	CC516004	CC516004 CH240_361
c	12	11.8	40.7	63	29	CG563472	CG563472 OST186777
c	13	11.8	40.7	65	9	AA733449	AA733449 vt73h08.r
	14	11.8	40.7	70	9	AI767928	AI767928 wi99c01.x
c	15	11.8	40.7	70	28	BH216023	BH216023 1006039G0
	16	11.6	40.0	49	14	U44334	U44334 ENU44334 As
c	17	11.6	40.0	58	9	AI584456	AI584456 fb93h12.x
c	18	11.6	40.0	65	29	CG519587	CG519587 OST83436
	19	11.6	40.0	70	9	AI814489	AI814489 wj73g11.x
c	20	11.4	39.3	32	29	HSMC42B09	X88068 H.sapiens D
	21	11.4	39.3	34	29	BX001854	BX001854 Arabidops
c	22	11.4	39.3	40	29	TA253H01Q	AL483109 T. brucei
	23	11.4	39.3	46	9	AI887082	AI887082 wl96e09.x
	24	11.4	39.3	54	29	CC556861	CC556861 CH240_464
	25	11.4	39.3	61	14	CD946166	CD946166 REJ_33 Ge
c	26	11.4	39.3	61	14	CD963134	CD963134 SDS_4 Gen
c	27	11.4	39.3	66	9	AA247859	AA247859 j337l.seq
	28	11.4	39.3	67	29	TA113E04Q	AL460350 T. brucei
	29	11.4	39.3	68	9	AU254479	AU254479 AU254479
	30	11.2	38.6	22	14	D18745	D18745 MUSGS01807
	31	11.2	38.6	30	9	AU259912	AU259912 AU259912
c	32	11.2	38.6	37	28	AZ950243	AZ950243 2M0214C15
	33	11.2	38.6	44	29	AL771575	AL771575 Arabidops
c	34	11.2	38.6	46	9	AA522160	AA522160 vf97c10.r
	35	11.2	38.6	49	9	AV841468	AV841468 AV841468
	36	11.2	38.6	54	28	BH224605	BH224605 1006120A0
	37	11.2	38.6	60	14	H19786	H19786 yn60a10.s1
	38	11.2	38.6	64	9	AI321110	AI321110 d4c09nm.r
c	39	11.2	38.6	65	29	AL755798	AL755798 Arabidops
	40	11.2	38.6	66	9	AI571487	AI571487 tr56g10.x
c	41	11.2	38.6	66	29	CG603111	CG603111 OST277306
c	42	11.2	38.6	67	9	AA936041	AA936041 nz53f10.s
	43	11.2	38.6	70	9	AU258214	AU258214 AU258214
c	44	11.2	38.6	70	14	CB916098	CB916098 VVD106F01
c	45	11	37.9	34	29	TA98E04P	AL461291 T. brucei

ALIGNMENTS

RESULT 1
AZ833686
LOCUS AZ833686 46 bp DNA linear GSS 20-FEB-2001
DEFINITION 2M0115L20R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0115L20 R, genomic survey sequence.
ACCESSION AZ833686
VERSION AZ833686.1 GI:13003594
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 46)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177

```
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0115  row: L  column: 20
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 46.
FEATURES              Location/Qualifiers
     source           1. .46
                      /organism="Mus musculus"
                      /mol_type="genomic DNA"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="UUGC2M0115L20"
                      /sex="Male"
                      /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                      /clone_lib="Mouse 10kb plasmid UUGC1M library"
                      /note="Vector: PWD42nv; Purified genomic DNA from M.
                      musculus C57BL/6J (male) was obtained  from the Jackson
                      Laboratory Mouse DNA Resource
                      (http://www.jax.org/resources/documents/dnares/). The DNA
                      was hydrodynamically sheared by repeated passage through a
                      0.005 inch orifice at constant velocity. The sheared DNA
                      was blunt end-repaired with T4 DNA polymerase and T4
                      polynucleotide kinase. Adaptor oligonucleotides were
                      ligated to the blunt ends in high  molar excess. The
                      adaptored DNA was purified and size-selected for a 9.5 to
                      10.5 kb range using preparative agarose gel
                      electrophoresis. Vector DNA was prepared from a derivative
                      of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                      inducible derivative of plasmid R1. The vector was ligated
                      with adaptors complementary to the insert adaptors and
                      purified. The sheared, adaptored mouse DNA was annealed to
                      adaptored vector DNA, and transformed into
                      chemically-competent E. coli XL10-Gold (Stratagene) cells
                      and selected for ampicillin resistance."
ORIGIN

  Query Match             44.8%;  Score 13;  DB 28;  Length 46;
  Best Local Similarity   50.0%;  Pred. No. 3.1e+04;
  Matches    8;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGC 19
              ||: |:::  |:||||
Db          8 GATACTTTAAGTAAGC 23


RESULT 2
BH759592
LOCUS       BH759592                  70 bp    DNA     linear   GSS 12-MAR-2002
DEFINITION  KG05236-3prime Drosophila melanogaster P{SUPor-P} P element
            insertion lines Drosophila melanogaster genomic Sequence recovered
            from 3' end of P element, genomic survey sequence.
ACCESSION   BH759592
VERSION     BH759592.1  GI:19352831
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 70)
  AUTHORS   Levis,R., Hoskins,R., Liao,G., Mozden,N., Tsang,G., He,Y.,
            Karpen,G., Bellen,H., Rubin,G. and Spradling,A.
  TITLE     The Berkeley Drosophila Genome Project Gene Disruption Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Gerald Rubin
            Berkeley Drosophila Genome Project
            University of California, Berkeley
            LSA Building, Berkeley, CA 94720-3200, USA
            Fax: 5106439947
            Email: gerry@fruitfly.berkeley.edu
            Sequence recovery method was inverse PCR.
             Sequence orientation is forward strand relative to 5' end of P
            element
             The P element insertion position is base 1 in the 70 bases. This
            insertion position refers to the first base of the 8 base target
            recognition sequence.
            Class: transposon-tagged.
FEATURES              Location/Qualifiers
     source           1. .70
                      /organism="Drosophila melanogaster"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:7227"
                      /clone_lib="Drosophila melanogaster P{SUPor-P} P element
                      insertion lines"
                      /note="Inverse PCR was performed on Drosophila
                      melanogaster strains each of which contains one or more
                      P{SUPor-P} P-element transposon insertion. The resultant
                      fragment for each strain was directly sequenced to
                      determine the genomic sequence at the site of insertion.
                      Details of the protocols used can be found at
                      http://www.fruitfly.org/about/methods/inverse.pcr.html."
ORIGIN

  Query Match             44.1%;  Score 12.8;  DB 28;  Length 70;
  Best Local Similarity   42.9%;  Pred. No. 4.1e+04;
  Matches    9;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          5 AUNCUUUNNGUAAGCCCNANG 25
              |: |:::   :|| ||| | |
Db         15 ATACTTTATTTAATCCCAAAG 35


RESULT 3
AA700959/c
LOCUS       AA700959                  52 bp    mRNA    linear   EST 19-DEC-1997
DEFINITION  zf87d10.s1 Soares_pineal_gland_N3HPG Homo sapiens cDNA clone
            IMAGE:383923 3' similar to TR:P79324 P79324 RIBOSOMAL PROTEIN L15
            ;, mRNA sequence.
ACCESSION   AA700959
VERSION     AA700959.1  GI:2704124
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 52)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
            Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
            Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .52
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="GDB:1292180"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:383923"
                      /lab_host="DH10B (ampicillin resistant)"
                      /clone_lib="Soares_pineal_gland_N3HPG"
                      /note="Organ: pineal gland; Vector: pT7T3D (Pharmacia)
                      with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
```

1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCCGTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaldo. "

ORIGIN

Query Match 42.8%; Score 12.4; DB 9; Length 52;
Best Local Similarity 52.9%; Pred. No. 6.4e+04;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```
Qy          4 GAUNCUUUNNGUAAGCC 20
              ||: | ::  |:|||||
Db         31 GATCCATTCCGTAAGCC 15
```

RESULT 4
AA468615/c
LOCUS AA468615 70 bp mRNA linear EST 15-AUG-1997
DEFINITION ne08b04.s1 NCI_CGAP_Co3 Homo sapiens cDNA clone IMAGE:880591 3' similar to TR:G178246 G178246 MATURE ALPHA-GALACTOSIDASE A, COMPLETE CDS PRECURSOR ;, mRNA sequence.
ACCESSION AA468615
VERSION AA468615.1 GI:2195149
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 70)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1285 Std Error: 0.00
Seq primer: -41m13 fwd. ET from Amersham
High quality sequence stop: 1.

FEATURES Location/Qualifiers
source 1. .70
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:880591"
/sex="pooled"
/tissue_type="colon"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Co3"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from 12 pooled bulk tumor samples and primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. "

ORIGIN

Query Match 42.8%; Score 12.4; DB 9; Length 70;
Best Local Similarity 52.9%; Pred. No. 6.6e+04;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```
Qy          4 GAUNCUUUNNGUAAGCC 20
              ||: |:::  | |||||
Db         64 GATCCTTTGGGCAAGCC 48
```

RESULT 5
AZ503560
LOCUS AZ503560 48 bp DNA linear GSS 05-OCT-2000
DEFINITION 1M0343E21F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0343E21 F, genomic survey sequence.
ACCESSION AZ503560
VERSION AZ503560.1 GI:10684876
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 48)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0343 row: E column: 21
Seq primer: CGTTGTAAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 48.

FEATURES Location/Qualifiers
source 1. .48
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0343E21"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 42.1%; Score 12.2; DB 28; Length 48;
Best Local Similarity 43.5%; Pred. No. 8e+04;
Matches 10; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

```
Qy          5 AUNCUUUNNGUAAGCCCNANGNG 27
              |: |: :  |:||| |  | | |
Db         13 ATGCTCTGGGTAAGGCACAAGAG 35


RESULT 6
BZ770420
LOCUS       BZ770420                 56 bp    DNA     linear   GSS 13-MAR-2003
DEFINITION  SALK_143355.56.00.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_143355.56.00.x, genomic
            survey sequence.
ACCESSION   BZ770420
VERSION     BZ770420.1  GI:28944104
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 56)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
            Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
            Shinn,P., Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA. This sequence lies within an annotated exon of At5g40030.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .56
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_143355.56.00.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
ORIGIN

  Query Match             42.1%;  Score 12.2;  DB 28;  Length 56;
  Best Local Similarity   45.5%;  Pred. No. 8.1e+04;
  Matches   10;  Conservative    4;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANG 25
              ||: |: :   :|||||  | |
Db          3 GATACTATTGTTAAGCCTAACG 24


RESULT 7
BX744082
LOCUS       BX744082                 66 bp    mRNA    linear   EST 18-NOV-2003
DEFINITION  BX744082 XGC-tadpole Silurana tropicalis cDNA clone TTpA072h19 3',
            mRNA sequence.
ACCESSION   BX744082
VERSION     BX744082.1  GI:38416822
KEYWORDS    EST.
SOURCE      Silurana tropicalis (western clawed frog)
  ORGANISM  Silurana tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Silurana.
REFERENCE   1  (bases 1 to 66)
  AUTHORS   Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
  TITLE     Sanger Xenopus tropicalis EST project 2001 (11_2003)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Croning MDR
            Sanger Institute
            Hinxton, Cambridgeshire, CB10 1SA, UK
            Email: trop@sanger.ac.uk
            Sanger Xenopus tropicalis EST project 2001
            TROPICALIS_SEQUENCE_ID: TTpA072h19.q1kaT7
            Sequencing primer: T7
            This sequence is from a Xenopus Gene Collection (XGC) library
            constructed by Nigel Garrett.
            cDNA was oligo dT primed from 5ug of poly A+ RNA from tadpole
            embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with
            EcoRI at the 5' end and NotI at the 3' end.
            Vector: pCS107; Site_1: EcoRI; Site_2: NotI
            Host: Escherichia coli DH10B.
FEATURES             Location/Qualifiers
     source          1. .66
                     /organism="Silurana tropicalis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8364"
                     /clone="TTpA072h19"
                     /dev_stage="tadpole (stage 35-40)"
                     /lab_host="E. coli DH10B"
                     /clone_lib="XGC-tadpole"
                     /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
                     was oligo dT primed from 5ug of poly A+ RNA from tadpole
                     embryos. EcoRI-NotI cut cDNA was then ligated into pCS107
                     with EcoRI at the 5' end and NotI at the 3' end"
ORIGIN

  Query Match             42.1%;  Score 12.2;  DB 13;  Length 66;
  Best Local Similarity   43.5%;  Pred. No. 8.3e+04;
  Matches   10;  Conservative    4;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          5 AUNCUUUNNGUAAGCCCNANGNG 27
              |: | ::   :|| ||| | | |
Db         20 ATGCCTTATTTAATCCCCATGTG 42


RESULT 8
CD946435/c
LOCUS       CD946435                 67 bp    mRNA    linear   EST 15-JUL-2003
DEFINITION  REN_47 GeneTag1 Zea mays cDNA, mRNA sequence.
ACCESSION   CD946435
VERSION     CD946435.1  GI:32794199
KEYWORDS    EST.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 67)
  AUTHORS   Genoplante.
  TITLE     Genoplante, a major partnership french program in plant genomics
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Genoplante
            Genoplante
            93, rue Henri Rochefort 91025 EVRY CEDEX France
            Tel: 33 1 69 47 54 00
            Fax: 33 1 69 47 54 10
            This sequence has been generated in the framework of the french
            plant genomics programme 'Genoplante' (http://www.genoplante.com
            and http://genoplante-info.infobiogen.fr).
FEATURES             Location/Qualifiers
     source          1. .67
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /cultivar="mixture"
                     /db_xref="taxon:4577"
```

```
                     /clone_lib="GeneTag1"
ORIGIN

  Query Match             42.1%;  Score 12.2;  DB 14;  Length 67;
  Best Local Similarity   50.0%;  Pred. No. 8.4e+04;
  Matches   11;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCCNANG 25
              ||: |: :  | | |||| | |
Db         25 GATACTCTGGGGATGCCCTAAG 4


RESULT 9
AA975071/c
LOCUS       AA975071                 40 bp    mRNA    linear   EST 26-AUG-1998
DEFINITION  on03d07.s1 NCI_CGAP_Kid3 Homo sapiens cDNA clone IMAGE:1555597 3'
            similar to TR:P70566 P70566 N-TROPOMODULIN. ;, mRNA sequence.
ACCESSION   AA975071
VERSION     AA975071.1  GI:3150863
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 2096   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1555597"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Kid3"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer,
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not
                     I and Eco RI sites of the modified pT7T3 vector. mRNA
                     source: 2 pooled kidneys. Library went through one round
                     of normalization. Library constructed by Bento Soares and
                     M. Fatima Bonaldo. "
ORIGIN

  Query Match             40.7%;  Score 11.8;  DB 9;  Length 40;
  Best Local Similarity   40.0%;  Pred. No. 1.2e+05;
  Matches    8;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          6 UNCUUUNNGUAAGCCCNANG 25
              : |::: |:||| ||   |
Db         35 TCCTTTCCGTAAGACCTTGG 16


RESULT 10
BE970036/c
LOCUS       BE970036                 49 bp    mRNA    linear   EST 04-OCT-2000
DEFINITION  601680150F1 NIH_MGC_78 Homo sapiens cDNA clone IMAGE:3950172 5',
            mRNA sequence.
ACCESSION   BE970036
VERSION     BE970036.1  GI:10582969
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONETECH Laboratories, Inc.
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM816  row: d  column: 13
            High quality sequence stop: 49.
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3950172"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_78"
                     /note="Organ: pancreas; Vector: pDNR-LIB (Clontech);
                     Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI
                     (ggccattatggcc); 5' and 3' adaptors were used in cloning
                     as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3'
                     and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C, G, or T). Average insert size 1.2
                     kb (range 0.5-4.0 kb). 14/15 colonies contained inserts
                     by PCR. This library was enriched for full-length clones
                     and was constructed by Clontech Laboratories (Palo Alto,
                     CA)."
ORIGIN

  Query Match             40.7%;  Score 11.8;  DB 10;  Length 49;
  Best Local Similarity   45.0%;  Pred. No. 1.3e+05;
  Matches    9;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          6 UNCUUUNNGUAAGCCCNANG 25
              :  :::  | ||||| | |
Db         48 TTTTTTATGCAAGCCCCAGG 29


RESULT 11
CC516004
LOCUS       CC516004                 51 bp    DNA     linear   GSS 17-JUN-2003
DEFINITION  CH240_361F9.T7 CHORI-240 Bos taurus genomic clone CH240_361F9,
            genomic survey sequence.
ACCESSION   CC516004
VERSION     CC516004.1  GI:31834292
KEYWORDS    GSS.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Holt,R., Stott,J., Yang,G., Barber,S., Smailus,D., Prabhu,A.-L.,
            Tsai,M., Cloutier,A., Lee,D., Girn,N., Olson,T., Mayo,M.,
            Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R.,
```

```
            Mathewson,C., Wye,N., Masson,A., Brown-John,M., Jones,S.,
            Schein,J., Marra,M., de Jong,P., McWilliam,S., Barris,W.,
            Dalrymple,B.P. and Tellam,R.
  TITLE     Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
  JOURNAL   Unpublished (2003)
COMMENT     Other_GSSs: CH240_361F9.TARBAC13P2
            Contact: Rob Holt
            Sequencing
            The British Columbia Cancer Agency Genome Science Centre
            600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
            Tel: 604-877-6085
            Fax: 604-877-6276
            Email: rholt@bcgsc.ca
            Clones are derived from the bovine BAC library CHORI-240
            (http://www.chori.org/bacpac/bovine240.htm). For BAC library
            availability, please contact Pieter de Jong (pdejong@mail.cho.org).
            Clones may be purchased from BACPAC Resources
            (http://www.chori.org/bacpac/ordering_information.htm). This work
            was undertaken as part of the International Bovine BAC Mapping
            Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
            British Columbia Genome Sciences Centre, Canada.
            Plate: 361  row: F  column: 9
            Seq primer: T7
            Class: BAC ends.
FEATURES                Location/Qualifiers
     source             1. .51
                        /organism="Bos taurus"
                        /mol_type="genomic DNA"
                        /strain="breed: Hereford"
                        /db_xref="taxon:9913"
                        /clone="CH240_361F9"
                        /sex="Male"
                        /cell_type="Blood"
                        /clone_lib="CHORI-240"
                        /note="Vector: pTARBAC1.3; Site_1: MboI; Site_2: MboI;
                        Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
                        library (Male) produced by Pieter de Jong"
ORIGIN

  Query Match             40.7%;  Score 11.8;  DB 29;  Length 51;
  Best Local Similarity   44.4%;  Pred. No. 1.3e+05;
  Matches    8;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCC 21
              ||:  :::  |: |||||
Db         31 GATGATTTCAGTGAGCCC 48


RESULT 12
CG563472/c
LOCUS       CG563472                63 bp    DNA     linear   GSS 01-OCT-2003
DEFINITION  OST186777 Mus musculus 129Sv/Ev Mus musculus genomic clone
            OST186777, genomic survey sequence.
ACCESSION   CG563472
VERSION     CG563472.1  GI:37350059
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 63)
  AUTHORS   Zambrowicz,B.P., Abuin,A., Ramirez-Solis,R., Richter,L.J.,
            Piggott,J., BeltrandelRio,H., Buxton,E.C., Edwards,J., Finch,R.A.,
            Friddle,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaing,C.,
            Key,B.W. Jr., Kipp,P., Kohlhauff,B., Ma,Z.-Q., Markesich,D.,
            Payne,R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,Z.-Z.,
            Sparks,M.J., Van Sligtenhorst,I., Vogel,P., Walke,W., Xu,N.,
            Zhu,Q., Person,C. and Sands,A.T.
  TITLE     Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap
            screen to identify potential targets for therapeutic intervention
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
COMMENT     Contact: Zambrowicz BP
            OmniBank
            Lexicon Genetics Incorporated
            4000 Research Forest Drive, The Woodlands, TX 77381, USA
            Email: materials@lexgen.com
            Gene trap sequence tag generated by 3' RACE from mouse ES cells as
            described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)
            Class: Gene Trap.
FEATURES                Location/Qualifiers
     source             1. .63
                        /organism="Mus musculus"
                        /mol_type="genomic DNA"
                        /strain="129SV/EV"
                        /db_xref="taxon:10090"
                        /clone="OST186777"
                        /cell_type="embryonic stem cell"
                        /clone_lib="Mus musculus 129Sv/Ev"
ORIGIN

  Query Match             40.7%;  Score 11.8;  DB 29;  Length 63;
  Best Local Similarity   50.0%;  Pred. No. 1.3e+05;
  Matches    9;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 GAUNCUUUNNGUAAGCCC 21
              | : |: :  |:||||||
Db         45 GTTTCTCTGTGTAAGCCC 28


RESULT 13
AA733449/c
LOCUS       AA733449                65 bp    mRNA    linear   EST 07-JAN-1998
DEFINITION  vt73h08.r1 Barstead mouse irradiated colon MPLRB7 Mus musculus cDNA
            clone IMAGE:1176831 5' similar to gb:X06617 40S RIBOSOMAL PROTEIN
            S11 (HUMAN);, mRNA sequence.
ACCESSION   AA733449
VERSION     AA733449.1  GI:2755116
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 65)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:634679
            Trace considered overall poor quality
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 1.
FEATURES                Location/Qualifiers
     source             1. .65
                        /organism="Mus musculus"
                        /mol_type="mRNA"
                        /strain="FVB/N"
                        /db_xref="taxon:10090"
                        /clone="IMAGE:1176831"
                        /dev_stage="8 weeks"
                        /lab_host="DH10B"
                        /clone_lib="Barstead mouse irradiated colon MPLRB7"
                        /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                        polylinker; Site_1: EcoRI; Site_2: NotI; Tissue obtained
                        from 8 week old mouse.  Colon was harvested 72 hours after
```

irradiation with 1400 Gys. 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors [AATTCGGATCCTTG], digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library constructed by Bob Barstead. "

ORIGIN

Query Match 40.7%; Score 11.8; DB 9; Length 65;
Best Local Similarity 44.4%; Pred. No. 1.3e+05;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

```
Qy          4 GAUNCUUUNNGUAAGCCC 21
              | : |:::  |:|||| |
Db         44 GCTGCTTTTGGTAAGCAC 27
```

RESULT 14
AI767928
LOCUS AI767928 70 bp mRNA linear EST 21-DEC-1999
DEFINITION wi99c01.x1 NCI_CGAP_Kid12 Homo sapiens cDNA clone IMAGE:2401440 3' similar to SW:RT14_HUMAN P78537 RT14 PROTEIN ;, mRNA sequence.
ACCESSION AI767928
VERSION AI767928.1 GI:5234426
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 70)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Insert Length: 574 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 1.

FEATURES Location/Qualifiers
source 1. .70
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2401440"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Kid12"
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP_Kid5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1323912-1325831, 1471368-1472903 and 1492104-1493255). Subtraction by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match 40.7%; Score 11.8; DB 9; Length 70;
Best Local Similarity 47.4%; Pred. No. 1.4e+05;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

```
Qy          5 AUNCUUUNNGUAAGCCCNA 23
              |: |:::  | ||||| |
Db         46 ATTCTTTAAGCAAGCCAGA 64
```

RESULT 15
BH216023/c
LOCUS BH216023 70 bp DNA linear GSS 08-NOV-2001
DEFINITION 1006039G04.2EL_y1 1006 - RescueMu Grid G Zea mays genomic, genomic survey sequence.
ACCESSION BH216023
VERSION BH216023.1 GI:16806681
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 70)
AUTHORS Walbot,V.
TITLE Maize genomic sequences found using engineered RescueMu transposon
JOURNAL Unpublished (2001)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Possible ligation site of ends cut by 2 different endonucleases.
Reverse complemented post-ligation sequence from source sequence.
Plate: 1006039 row: 43
Class: transposon-tagged.

FEATURES Location/Qualifiers
source 1. .70
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="1006 - RescueMu Grid G"
/note="Organ: leaf; Vector: RescueMu (engineered from pBlueScript backbone); Site_1: BamHI; Site_2: BglII; RescueMu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site 'www.zmdb.iastate.edu' and follow the links for 'RescueMu.' Grid G was grown at Stanford in 2000. DNA was extracted from leaf punches, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

ORIGIN

Query Match 40.7%; Score 11.8; DB 28; Length 70;
Best Local Similarity 50.0%; Pred. No. 1.4e+05;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```
Qy          4 GAUNCUUUNNGUAAGCCC 21
              ||: |:::  | | ||||
Db         70 GATCCTTTTAGGAGGCCC 53
```

Search completed: March 23, 2004, 17:05:35
Job time : 2020.33 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 11:34:35 ; Search time 235.333 Seconds
(without alignments)
523.503 Million cell updates/sec

Title: US-09-310-844C-24
Perfect score: 29
Sequence: 1 uaugauucuuuuuguaagcccuaggggcu 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 3369620

Minimum DB seq length: 0
Maximum DB seq length: 70

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	100.0	29	3	AAA70828	Aaa70828 Molecular
2	29	100.0	42	3	AAA71123	Aaa71123 Molecular
3	29	100.0	42	3	AAA71131	Aaa71131 Molecular
4	28	96.6	45	3	AAA70824	Aaa70824 Molecular
5	28	96.6	46	3	AAA71087	Aaa71087 Molecular
6	28	96.6	46	3	AAA71096	Aaa71096 Molecular
7	28	96.6	46	3	AAA71099	Aaa71099 Molecular
8	28	96.6	46	3	AAA71100	Aaa71100 Molecular
9	28	96.6	46	3	AAA71104	Aaa71104 Molecular
10	25.8	89.0	42	3	AAA71113	Aaa71113 Molecular
11	25.8	89.0	42	3	AAA71118	Aaa71118 Molecular
12	25.8	89.0	42	3	AAA71126	Aaa71126 Molecular
13	24.8	85.5	46	3	AAA71085	Aaa71085 Molecular
14	24.8	85.5	46	3	AAA71103	Aaa71103 Molecular
15	23.8	82.1	42	3	AAA71114	Aaa71114 Molecular
16	23.8	82.1	42	3	AAA71119	Aaa71119 Molecular
17	23.8	82.1	42	3	AAA71127	Aaa71127 Molecular
18	23.8	82.1	46	3	AAA71094	Aaa71094 Molecular
19	23.8	82.1	46	3	AAA71110	Aaa71110 Molecular
20	23.2	80.0	29	3	AAA70829	Aaa70829 Molecular
21	23.2	80.0	29	3	AAA70830	Aaa70830 Molecular
22	23.2	80.0	42	3	AAA71121	Aaa71121 Molecular
23	23.2	80.0	42	3	AAA71128	Aaa71128 Molecular
24	23.2	80.0	42	3	AAA71120	Aaa71120 Molecular
25	23.2	80.0	42	3	AAA71116	Aaa71116 Molecular
26	23.2	80.0	42	3	AAA71115	Aaa71115 Molecular
27	23.2	80.0	42	3	AAA71129	Aaa71129 Molecular
28	22.6	77.9	42	3	AAA71124	Aaa71124 Molecular
29	22.6	77.9	42	3	AAA71132	Aaa71132 Molecular
30	22.2	76.6	45	3	AAA70826	Aaa70826 Molecular
31	22.2	76.6	45	3	AAA70825	Aaa70825 Molecular
32	22.2	76.6	46	3	AAA71089	Aaa71089 Molecular
33	22.2	76.6	46	3	AAA71106	Aaa71106 Molecular
34	22.2	76.6	46	3	AAA71107	Aaa71107 Molecular
35	22.2	76.6	46	3	AAA71088	Aaa71088 Molecular
36	22.2	76.6	46	3	AAA71105	Aaa71105 Molecular
37	22.2	76.6	46	3	AAA71090	Aaa71090 Molecular
38	21.6	74.5	46	3	AAA71111	Aaa71111 Molecular
39	21.6	74.5	46	3	AAA71095	Aaa71095 Molecular
40	21.6	74.5	46	3	AAA71109	Aaa71109 Molecular
41	21.6	74.5	46	3	AAA71093	Aaa71093 Molecular
42	19.4	66.9	46	3	AAA71098	Aaa71098 Molecular
43	19.4	66.9	46	3	AAA71102	Aaa71102 Molecular
44	19.4	66.9	46	3	AAA71084	Aaa71084 Molecular
45	18.4	63.4	42	3	AAA71130	Aaa71130 Molecular

ALIGNMENTS

RESULT 1
AAA70828
ID AAA70828 standard; RNA; 29 BP.
XX
AC AAA70828;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #28.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Homo sapiens.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Claim 235; Page 235; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The

CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 29 BP; 5 A; 5 C; 7 G; 0 T; 12 U; 0 Other;

Query Match 100.0%; Score 29; DB 3; Length 29;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 UAUGAUUCUUUUUGUAAGCCCUAGGGGCU 29
||||||||||||||||||||||||||||||
Db 1 UAUGAUUCUUUUUGUAAGCCCUAGGGGCU 29

RESULT 2
AAA71123
ID AAA71123 standard; DNA; 42 BP.
XX
AC AAA71123;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #129.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 125; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 42 BP; 9 A; 6 C; 9 G; 18 T; 0 U; 0 Other;

Query Match 100.0%; Score 29; DB 3; Length 42;
Best Local Similarity 58.6%; Pred. No. 0.0027;
Matches 17; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

Qy 1 UAUGAUUCUUUUUGUAAGCCCUAGGGGCU 29
:|:||::|:::::|:||||||:||||||:
Db 4 TATGATTCTTTTTGTAAGCCCTAGGGGCT 32

RESULT 3
AAA71131
ID AAA71131 standard; RNA; 42 BP.
XX
AC AAA71131;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #200.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 126; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary

```
CC   structure defined by: (a) 3 nucleotides forming a first side of a first
CC   double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC   internal loop region; (c) 4 nucleotides forming a first side of a second
CC   ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC   nucleotides forming a second side of the second ds region; (f) 4
CC   nucleotides forming a second side of the internal loop region; and (g) 3
CC   nucleotides forming a second side of the first ds region; (2) a purified
CC   and isolated RNA fragment comprising the human sequence
CC   UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC   used for identifying agents which modulate the activity of biomolecules,
CC   particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC   or industrial compounds
XX
SQ   Sequence 42 BP; 9 A; 6 C; 9 G; 0 T; 18 U; 0 Other;

  Query Match             100.0%;  Score 29;  DB 3;  Length 42;
  Best Local Similarity   100.0%;  Pred. No. 0.0027;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 UAUGAUUCUUUUUGUAAGCCCUAGGGGCU 29
              |||||||||||||||||||||||||||||
Db          4 UAUGAUUCUUUUUGUAAGCCCUAGGGGCU 32


RESULT 4
AAA70824
ID   AAA70824 standard; RNA; 45 BP.
XX
AC   AAA70824;
XX
DT   27-APR-2001  (first entry)
XX
DE   Molecular interaction site RNA #24.
XX
KW   Modulator; identification; molecular interaction; virtual library; ss.
XX
OS   Homo sapiens.
XX
PN   WO9958947-A2.
XX
PD   18-NOV-1999.
XX
PF   12-MAY-1999;   99WO-US010361.
XX
PR   12-MAY-1998;   98US-00076404.
PR   12-MAY-1998;   98US-0085092P.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Ecker DJ,  Griffey R,  Crooke ST,  Sampath R,  Swayze E,  Mohan V;
PI   Hofstadler S,  Mcneil J;
XX
DR   WPI; 2000-086439/07.
XX
PT   Identifying compounds which modulate activity of target biomolecules,
PT   used to provide compounds which can be used as pharmacological,
PT   agricultural and industrial compounds.
XX
PS   Claim 220; Page 232; 405pp; English.
XX
CC   This invention describes a novel method for identifying compounds which
CC   modulate the activity of a target biomolecule. The method uses 3-
CC   dimensional representations of the biomolecule and a library of compounds
CC   and comprises (a) identifying at least one molecular interaction site of
CC   the target RNA; (b) generating in silico a virtual library of compounds
CC   predicted or calculated to interact with the molecular interaction site;
CC   and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC   with members of the virtual library of compounds to generate a hierarchy
CC   of the compounds ranked in accordance with their respective ability to
CC   form physical interactions with the molecular interaction site. The
CC   method also describes (1) RNA comprising a joined sequence of at least 24
CC   nucleotides but not more than 70 nucleotides and having secondary
CC   structure defined by: (a) 3 nucleotides forming a first side of a first
CC   double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC   internal loop region; (c) 4 nucleotides forming a first side of a second
CC   ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC   nucleotides forming a second side of the second ds region; (f) 4
CC   nucleotides forming a second side of the internal loop region; and (g) 3
CC   nucleotides forming a second side of the first ds region; (2) a purified
CC   and isolated RNA fragment comprising the human sequence
CC   UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC   used for identifying agents which modulate the activity of biomolecules,
CC   particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC   or industrial compounds
XX
SQ   Sequence 45 BP; 11 A; 6 C; 9 G; 0 T; 19 U; 0 Other;

  Query Match             96.6%;  Score 28;  DB 3;  Length 45;
  Best Local Similarity   100.0%;  Pred. No. 0.0077;
  Matches   28;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 UAUGAUUCUUUUUGUAAGCCCUAGGGGC 28
              ||||||||||||||||||||||||||||
Db         18 UAUGAUUCUUUUUGUAAGCCCUAGGGGC 45


RESULT 5
AAA71087
ID   AAA71087 standard; DNA; 46 BP.
XX
AC   AAA71087;
XX
DT   27-APR-2001  (first entry)
XX
DE   Molecular interaction site DNA #110.
XX
KW   Modulator; identification; molecular interaction; virtual library; ss.
XX
OS   Unidentified.
XX
PN   WO9958947-A2.
XX
PD   18-NOV-1999.
XX
PF   12-MAY-1999;   99WO-US010361.
XX
PR   12-MAY-1998;   98US-00076404.
PR   12-MAY-1998;   98US-0085092P.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Ecker DJ,  Griffey R,  Crooke ST,  Sampath R,  Swayze E,  Mohan V;
PI   Hofstadler S,  Mcneil J;
XX
DR   WPI; 2000-086439/07.
XX
PT   Identifying compounds which modulate activity of target biomolecules,
PT   used to provide compounds which can be used as pharmacological,
PT   agricultural and industrial compounds.
XX
PS   Example 7; Fig 121; 405pp; English.
XX
CC   This invention describes a novel method for identifying compounds which
CC   modulate the activity of a target biomolecule. The method uses 3-
CC   dimensional representations of the biomolecule and a library of compounds
CC   and comprises (a) identifying at least one molecular interaction site of
CC   the target RNA; (b) generating in silico a virtual library of compounds
CC   predicted or calculated to interact with the molecular interaction site;
CC   and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC   with members of the virtual library of compounds to generate a hierarchy
CC   of the compounds ranked in accordance with their respective ability to
CC   form physical interactions with the molecular interaction site. The
CC   method also describes (1) RNA comprising a joined sequence of at least 24
CC   nucleotides but not more than 70 nucleotides and having secondary
CC   structure defined by: (a) 3 nucleotides forming a first side of a first
CC   double stranded (ds) region; (b) 2 nucleotides forming a first side of an
```

CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 46 BP; 11 A; 7 C; 9 G; 19 T; 0 U; 0 Other;

Query Match 96.6%; Score 28; DB 3; Length 46;
Best Local Similarity 60.7%; Pred. No. 0.0077;
Matches 17; Conservative 11; Mismatches 0; Indels 0; Gaps 0;

Qy 1 UAUGAUUCUUUUUGUAAGCCCUAGGGGC 28
:|:||::|:::::|:||||||:||||||
Db 19 TATGATTCTTTTTGTAAGCCCTAGGGGC 46

RESULT 6
AAA71096
ID AAA71096 standard; DNA; 46 BP.
XX
AC AAA71096;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #119.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 121; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 46 BP; 11 A; 7 C; 9 G; 19 T; 0 U; 0 Other;

Query Match 96.6%; Score 28; DB 3; Length 46;
Best Local Similarity 60.7%; Pred. No. 0.0077;
Matches 17; Conservative 11; Mismatches 0; Indels 0; Gaps 0;

Qy 1 UAUGAUUCUUUUUGUAAGCCCUAGGGGC 28
:|:||::|:::::|:||||||:||||||
Db 19 TATGATTCTTTTTGTAAGCCCTAGGGGC 46

RESULT 7
AAA71099
ID AAA71099 standard; DNA; 46 BP.
XX
AC AAA71099;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #122.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 121; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4

CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 46 BP; 11 A; 7 C; 9 G; 19 T; 0 U; 0 Other;

Query Match 96.6%; Score 28; DB 3; Length 46;
Best Local Similarity 60.7%; Pred. No. 0.0077;
Matches 17; Conservative 11; Mismatches 0; Indels 0; Gaps 0;

Qy 1 UAUGAUUCUUUUUGUAAGCCCUAGGGGC 28
:|:||::|:::::|:||||||:||||||
Db 19 TATGATTCTTTTTGTAAGCCCTAGGGGC 46

RESULT 8
AAA71100
ID AAA71100 standard; DNA; 46 BP.
XX
AC AAA71100;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #123.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 121; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 46 BP; 11 A; 7 C; 9 G; 19 T; 0 U; 0 Other;

Query Match 96.6%; Score 28; DB 3; Length 46;
Best Local Similarity 60.7%; Pred. No. 0.0077;
Matches 17; Conservative 11; Mismatches 0; Indels 0; Gaps 0;

Qy 1 UAUGAUUCUUUUUGUAAGCCCUAGGGGC 28
:|:||::|:::::|:||||||:||||||
Db 19 TATGATTCTTTTTGTAAGCCCTAGGGGC 46

RESULT 9
AAA71104
ID AAA71104 standard; RNA; 46 BP.
XX
AC AAA71104;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #180.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 122; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3

CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 46 BP; 11 A; 7 C; 9 G; 0 T; 19 U; 0 Other;

Query Match 96.6%; Score 28; DB 3; Length 46;
Best Local Similarity 100.0%; Pred. No. 0.0077;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 UAUGAUUCUUUUUGUAAGCCCUAGGGGC 28
||||||||||||||||||||||||||||
Db 19 UAUGAUUCUUUUUGUAAGCCCUAGGGGC 46

RESULT 10
AAA71113
ID AAA71113 standard; RNA; 42 BP.
XX
AC AAA71113;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #189.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 122; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 42 BP; 12 A; 7 C; 6 G; 0 T; 17 U; 0 Other;

Query Match 89.0%; Score 25.8; DB 3; Length 42;
Best Local Similarity 93.1%; Pred. No. 0.074;
Matches 27; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 UAUGAUUCUUUUUGUAAGCCCUAGGGGCU 29
|| |||||||||||||||||||| |||||
Db 4 UAAGAUUCUUUUUGUAAGCCCUACGGGCU 32

RESULT 11
AAA71118
ID AAA71118 standard; DNA; 42 BP.
XX
AC AAA71118;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #124.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 125; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence

CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 42 BP; 12 A; 7 C; 6 G; 17 T; 0 U; 0 Other;

Query Match 89.0%; Score 25.8; DB 3; Length 42;
Best Local Similarity 55.2%; Pred. No. 0.074;
Matches 16; Conservative 11; Mismatches 2; Indels 0; Gaps 0;

Qy 1 UAUGAUUCUUUUUGUAAGCCCUAGGGGCU 29
:| ||::|:::::|:||||||:| ||||:
Db 4 TAAGATTCTTTTTGTAAGCCCTACGGGCT 32

RESULT 12
AAA71126
ID AAA71126 standard; RNA; 42 BP.
XX
AC AAA71126;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #195.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 126; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 42 BP; 12 A; 7 C; 6 G; 0 T; 17 U; 0 Other;

Query Match 89.0%; Score 25.8; DB 3; Length 42;
Best Local Similarity 93.1%; Pred. No. 0.074;
Matches 27; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 UAUGAUUCUUUUUGUAAGCCCUAGGGGCU 29
|| |||||||||||||||||||| |||||
Db 4 UAAGAUUCUUUUUGUAAGCCCUACGGGCU 32

RESULT 13
AAA71085
ID AAA71085 standard; DNA; 46 BP.
XX
AC AAA71085;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #108.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 121; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,

CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 46 BP; 12 A; 7 C; 9 G; 18 T; 0 U; 0 Other;

Query Match 85.5%; Score 24.8; DB 3; Length 46;
Best Local Similarity 57.1%; Pred. No. 0.21;
Matches 16; Conservative 10; Mismatches 2; Indels 0; Gaps 0;

Qy 1 UAUGAUUCUUUUUGUAAGCCCUAGGGGC 28
:| ||::|:::::|:||||||:| ||||
Db 19 TAAGATTCTTTTTGTAAGCCCTACGGGC 46

RESULT 14
AAA71103
ID AAA71103 standard; RNA; 46 BP.
XX
AC AAA71103;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #179.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 122; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 46 BP; 12 A; 7 C; 9 G; 0 T; 18 U; 0 Other;

Query Match 85.5%; Score 24.8; DB 3; Length 46;
Best Local Similarity 92.9%; Pred. No. 0.21;
Matches 26; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 UAUGAUUCUUUUUGUAAGCCCUAGGGGC 28
|| |||||||||||||||||||| ||||
Db 19 UAAGAUUCUUUUUGUAAGCCCUACGGGC 46

RESULT 15
AAA71114
ID AAA71114 standard; RNA; 42 BP.
XX
AC AAA71114;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #190.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 122; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds

XX
SQ Sequence 42 BP; 11 A; 8 C; 7 G; 0 T; 16 U; 0 Other;

Query Match 82.1%; Score 23.8; DB 3; Length 42;
Best Local Similarity 92.6%; Pred. No. 0.58;
Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 UAUGAUUCUUUUUGUAAGCCCUAGGGG 27
              || |||||||||||||||||||||||| |
Db          4 UAAGAUUCUUUUUGUAAGCCCUAGGCG 30
```

Search completed: March 23, 2004, 14:53:14
Job time : 235.333 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 11:44:09 ; Search time 633.333 Seconds
(without alignments)
1984.655 Million cell updates/sec

Title: US-09-310-844C-24
Perfect score: 29
Sequence: 1 uaugauucuuuuuguaagcccuaggggcu 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 1733942

Minimum DB seq length: 0
Maximum DB seq length: 70

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	100.0	42	6	BD274275	BD274275 Identific
2	29	100.0	42	6	BD274283	BD274283 Identific
3	28	96.6	46	6	BD274240	BD274240 Identific
4	28	96.6	46	6	BD274249	BD274249 Identific
5	28	96.6	46	6	BD274252	BD274252 Identific
6	28	96.6	46	6	BD274253	BD274253 Identific
7	28	96.6	46	6	BD274257	BD274257 Identific
8	28	96.6	46	6	BD274265	BD274265 Identific
9	28	96.6	46	6	BD274268	BD274268 Identific
10	28	96.6	46	6	BD274269	BD274269 Identific
11	25.8	89.0	42	6	BD274270	BD274270 Identific
12	25.8	89.0	42	6	BD274278	BD274278 Identific
13	24.8	85.5	46	6	BD274238	BD274238 Identific
14	24.8	85.5	46	6	BD274256	BD274256 Identific
15	23.8	82.1	42	6	BD274271	BD274271 Identific
16	23.8	82.1	42	6	BD274279	BD274279 Identific
17	23.8	82.1	46	6	BD274247	BD274247 Identific
18	23.8	82.1	46	6	BD274263	BD274263 Identific
19	23.2	80.0	42	6	BD274272	BD274272 Identific
20	23.2	80.0	42	6	BD274273	BD274273 Identific
21	23.2	80.0	42	6	BD274280	BD274280 Identific
22	23.2	80.0	42	6	BD274281	BD274281 Identific
23	22.6	77.9	42	6	BD274284	BD274284 Identific
24	22.2	76.6	46	6	BD274241	BD274241 Identific
25	22.2	76.6	46	6	BD274242	BD274242 Identific
26	22.2	76.6	46	6	BD274243	BD274243 Identific
27	22.2	76.6	46	6	BD274258	BD274258 Identific
28	22.2	76.6	46	6	BD274259	BD274259 Identific
29	22.2	76.6	46	6	BD274260	BD274260 Identific
30	21.6	74.5	46	6	BD274246	BD274246 Identific
31	21.6	74.5	46	6	BD274248	BD274248 Identific
32	21.6	74.5	46	6	BD274262	BD274262 Identific
33	21.6	74.5	46	6	BD274264	BD274264 Identific
34	19.4	66.9	42	6	BD274276	BD274276 Identific
35	19.4	66.9	46	6	BD274237	BD274237 Identific
36	19.4	66.9	46	6	BD274251	BD274251 Identific
37	19.4	66.9	46	6	BD274255	BD274255 Identific
38	19.4	66.9	46	6	BD274267	BD274267 Identific
39	18.4	63.4	42	6	BD274274	BD274274 Identific
40	18.4	63.4	42	6	BD274282	BD274282 Identific
41	18	62.1	44	6	BD274277	BD274277 Identific
c 42	15.8	54.5	30	6	AX792238	AX792238 Sequence
c 43	15.4	53.1	41	6	AX514720	AX514720 Sequence
c 44	15.4	53.1	41	6	AX520728	AX520728 Sequence
c 45	15.2	52.4	33	6	AR020509	AR020509 Sequence

ALIGNMENTS

RESULT 1
BD274275
LOCUS BD274275 42 bp DNA linear PAT 17-JUL-2003
DEFINITION Identification of molecular interaction sites in RNA for novel drug discovery.
ACCESSION BD274275
VERSION BD274275.1 GI:33084043
KEYWORDS JP 2002526030-A/242.
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 42)
AUTHORS Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
TITLE Identification of molecular interaction sites in RNA for novel drug discovery

```
  JOURNAL   Patent: JP 2002526030-A 242 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/242
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .42
            FT                   /organism='Artificial Sequence'.
FEATURES              Location/Qualifiers
     source           1. .42
                      /organism="synthetic construct"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:32630"
ORIGIN

  Query Match             100.0%;  Score 29;  DB 6;  Length 42;
  Best Local Similarity   58.6%;  Pred. No. 0.021;
  Matches   17;  Conservative   12;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 UAUGAUUCUUUUUGUAAGCCCUAGGGGCU 29
              :|:||::|:::::|:||||||:||||||:
Db          4 TATGATTCTTTTTGTAAGCCCTAGGGGCT 32


RESULT 2
BD274283
LOCUS       BD274283                 42 bp    RNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274283
VERSION     BD274283.1  GI:33084051
KEYWORDS    JP 2002526030-A/250.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 250 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/250
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .42
            FT                   /organism='Artificial Sequence'.
FEATURES              Location/Qualifiers
     source           1. .42
                      /organism="synthetic construct"
                      /mol_type="genomic RNA"
                      /db_xref="taxon:32630"
ORIGIN

  Query Match             100.0%;  Score 29;  DB 6;  Length 42;
  Best Local Similarity   58.6%;  Pred. No. 0.021;
  Matches   17;  Conservative   12;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 UAUGAUUCUUUUUGUAAGCCCUAGGGGCU 29
              :|:||::|:::::|:||||||:||||||:
Db          4 TATGATTCTTTTTGTAAGCCCTAGGGGCT 32


RESULT 3
BD274240
LOCUS       BD274240                 46 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274240
VERSION     BD274240.1  GI:33084008
KEYWORDS    JP 2002526030-A/207.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 207 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/207
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .46
            FT                   /organism='Artificial Sequence'.
FEATURES              Location/Qualifiers
     source           1. .46
                      /organism="synthetic construct"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:32630"
ORIGIN

  Query Match             96.6%;  Score 28;  DB 6;  Length 46;
  Best Local Similarity   60.7%;  Pred. No. 0.061;
  Matches   17;  Conservative   11;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 UAUGAUUCUUUUUGUAAGCCCUAGGGGC 28
              :|:||::|:::::|:||||||:|||||
Db         19 TATGATTCTTTTTGTAAGCCCTAGGGGC 46


RESULT 4
BD274249
LOCUS       BD274249                 46 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274249
VERSION     BD274249.1  GI:33084017
KEYWORDS    JP 2002526030-A/216.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 216 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/216
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .46
            FT                   /organism='Artificial Sequence'.
FEATURES              Location/Qualifiers
```

```
     source          1. .46
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             96.6%;  Score 28;  DB 6;  Length 46;
  Best Local Similarity   60.7%;  Pred. No. 0.061;
  Matches   17;  Conservative   11;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 UAUGAUUCUUUUUGUAAGCCCUAGGGGC 28
              :|:||::|:::::|:||||||:||||||
Db         19 TATGATTCTTTTTGTAAGCCCTAGGGGC 46


RESULT 5
BD274252
LOCUS       BD274252                  46 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274252
VERSION     BD274252.1  GI:33084020
KEYWORDS    JP 2002526030-A/219.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 219 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/219
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .46
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             96.6%;  Score 28;  DB 6;  Length 46;
  Best Local Similarity   60.7%;  Pred. No. 0.061;
  Matches   17;  Conservative   11;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 UAUGAUUCUUUUUGUAAGCCCUAGGGGC 28
              :|:||::|:::::|:||||||:||||||
Db         19 TATGATTCTTTTTGTAAGCCCTAGGGGC 46


RESULT 6
BD274253
LOCUS       BD274253                  46 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274253
VERSION     BD274253.1  GI:33084021
KEYWORDS    JP 2002526030-A/220.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 220 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/220
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .46
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             96.6%;  Score 28;  DB 6;  Length 46;
  Best Local Similarity   60.7%;  Pred. No. 0.061;
  Matches   17;  Conservative   11;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 UAUGAUUCUUUUUGUAAGCCCUAGGGGC 28
              :|:||::|:::::|:||||||:||||||
Db         19 TATGATTCTTTTTGTAAGCCCTAGGGGC 46


RESULT 7
BD274257
LOCUS       BD274257                  46 bp    RNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274257
VERSION     BD274257.1  GI:33084025
KEYWORDS    JP 2002526030-A/224.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 224 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/224
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .46
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="synthetic construct"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             96.6%;  Score 28;  DB 6;  Length 46;
  Best Local Similarity   60.7%;  Pred. No. 0.061;
  Matches   17;  Conservative   11;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 UAUGAUUCUUUUUGUAAGCCCUAGGGGC 28
              :|:||::|:::::|:||||||:||||||
Db         19 TATGATTCTTTTTGTAAGCCCTAGGGGC 46
```

```
RESULT 8
BD274265
LOCUS       BD274265                 46 bp    RNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274265
VERSION     BD274265.1  GI:33084033
KEYWORDS    JP 2002526030-A/232.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 232 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/232
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .46
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="synthetic construct"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             96.6%;  Score 28;  DB 6;  Length 46;
  Best Local Similarity   60.7%;  Pred. No. 0.061;
  Matches   17;  Conservative   11;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 UAUGAUUCUUUUUGUAAGCCCUAGGGGC 28
              :|:||::|:::::|:||||||:||||||
Db         19 TATGATTCTTTTTGTAAGCCCTAGGGGC 46


RESULT 9
BD274268
LOCUS       BD274268                 46 bp    RNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274268
VERSION     BD274268.1  GI:33084036
KEYWORDS    JP 2002526030-A/235.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 235 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/235
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .46
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="synthetic construct"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             96.6%;  Score 28;  DB 6;  Length 46;
  Best Local Similarity   60.7%;  Pred. No. 0.061;
  Matches   17;  Conservative   11;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 UAUGAUUCUUUUUGUAAGCCCUAGGGGC 28
              :|:||::|:::::|:||||||:||||||
Db         19 TATGATTCTTTTTGTAAGCCCTAGGGGC 46


RESULT 10
BD274269
LOCUS       BD274269                 46 bp    RNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274269
VERSION     BD274269.1  GI:33084037
KEYWORDS    JP 2002526030-A/236.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 236 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/236
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .46
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="synthetic construct"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             96.6%;  Score 28;  DB 6;  Length 46;
  Best Local Similarity   60.7%;  Pred. No. 0.061;
  Matches   17;  Conservative   11;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 UAUGAUUCUUUUUGUAAGCCCUAGGGGC 28
              :|:||::|:::::|:||||||:||||||
Db         19 TATGATTCTTTTTGTAAGCCCTAGGGGC 46


RESULT 11
BD274270
LOCUS       BD274270                 42 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274270
VERSION     BD274270.1  GI:33084038
KEYWORDS    JP 2002526030-A/237.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
```

```
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 237 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/237
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .42
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .42
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             89.0%;  Score 25.8;  DB 6;  Length 42;
  Best Local Similarity   55.2%;  Pred. No. 0.63;
  Matches   16;  Conservative   11;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 UAUGAUUCUUUUUGUAAGCCCUAGGGGCU 29
              :| ||::|:::::|:||||||:| ||||:
Db          4 TAAGATTCTTTTTGTAAGCCCTACGGGCT 32


RESULT 12
BD274278
LOCUS       BD274278                  42 bp    RNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274278
VERSION     BD274278.1  GI:33084046
KEYWORDS    JP 2002526030-A/245.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 245 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/245
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .42
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .42
                     /organism="synthetic construct"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             89.0%;  Score 25.8;  DB 6;  Length 42;
  Best Local Similarity   55.2%;  Pred. No. 0.63;
  Matches   16;  Conservative   11;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 UAUGAUUCUUUUUGUAAGCCCUAGGGGCU 29
              :| ||::|:::::|:||||||:| ||||:
Db          4 TAAGATTCTTTTTGTAAGCCCTACGGGCT 32


RESULT 13
BD274238
LOCUS       BD274238                  46 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274238
VERSION     BD274238.1  GI:33084006
KEYWORDS    JP 2002526030-A/205.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 205 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/205
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .46
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             85.5%;  Score 24.8;  DB 6;  Length 46;
  Best Local Similarity   57.1%;  Pred. No. 1.8;
  Matches   16;  Conservative   10;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 UAUGAUUCUUUUUGUAAGCCCUAGGGGC 28
              :| ||::|:::::|:||||||:| ||||
Db         19 TAAGATTCTTTTTGTAAGCCCTACGGGC 46


RESULT 14
BD274256
LOCUS       BD274256                  46 bp    RNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274256
VERSION     BD274256.1  GI:33084024
KEYWORDS    JP 2002526030-A/223.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 223 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/223
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .46
```

```
           FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="synthetic construct"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             85.5%;  Score 24.8;  DB 6;  Length 46;
  Best Local Similarity   57.1%;  Pred. No. 1.8;
  Matches   16;  Conservative   10;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 UAUGAUUCUUUUUGUAAGCCCUAGGGGC 28
              :| ||::|:::::|:||||||:| ||||
Db         19 TAAGATTCTTTTTGTAAGCCCTACGGGC 46


RESULT 15
BD274271
LOCUS       BD274271                 42 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274271
VERSION     BD274271.1  GI:33084039
KEYWORDS    JP 2002526030-A/238.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 238 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/238
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .42
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .42
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             82.1%;  Score 23.8;  DB 6;  Length 42;
  Best Local Similarity   55.6%;  Pred. No. 5.3;
  Matches   15;  Conservative   10;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 UAUGAUUCUUUUUGUAAGCCCUAGGGG 27
              :| ||::|:::::|:||||||:||| |
Db          4 TAAGATTCTTTTTGTAAGCCCTAGGCG 30


Search completed: March 23, 2004, 15:25:08
Job time : 634.333 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 14:41:54 ; Search time 55.3333 Seconds
(without alignments)
290.848 Million cell updates/sec

Title: US-09-310-844C-24
Perfect score: 29
Sequence: 1 uaugauucuuuuuguaagcccuaggggcu 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 895828

Minimum DB seq length: 0
Maximum DB seq length: 70

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	15.2	52.4	25	3	US-08-943-731-336	Sequence 336, App
c	2	15.2	52.4	33	1	US-08-667-079B-5	Sequence 5, Appli
c	3	14.8	51.0	35	6	5422260-12	Patent No. 5422260
c	4	14.8	51.0	36	3	US-09-440-001-1	Sequence 1, Appli
c	5	14.8	51.0	36	4	US-09-605-685-1	Sequence 1, Appli
c	6	14.2	49.0	25	4	US-09-827-998-1098	Sequence 1098, Ap
c	7	14.2	49.0	25	4	US-09-827-998-1099	Sequence 1099, Ap
c	8	14.2	49.0	25	4	US-09-827-998-1100	Sequence 1100, Ap
c	9	14.2	49.0	25	4	US-09-827-998-1101	Sequence 1101, Ap
c	10	14.2	49.0	25	4	US-09-827-998-1102	Sequence 1102, Ap
c	11	14.2	49.0	25	4	US-09-827-998-1103	Sequence 1103, Ap
c	12	14.2	49.0	25	4	US-09-827-998-1104	Sequence 1104, Ap
	13	14	48.3	30	4	US-09-690-146A-5	Sequence 5, Appli
c	14	14	48.3	30	4	US-09-690-146A-7	Sequence 7, Appli
	15	14	48.3	37	1	US-08-049-264C-55	Sequence 55, Appl
	16	14	48.3	37	1	US-08-476-562-55	Sequence 55, Appl
	17	14	48.3	37	1	US-08-479-723A-55	Sequence 55, Appl
	18	14	48.3	37	5	PCT-US94-04310-55	Sequence 55, Appl
c	19	14	48.3	44	1	US-08-049-264C-54	Sequence 54, Appl
c	20	14	48.3	44	1	US-08-476-562-54	Sequence 54, Appl
c	21	14	48.3	44	1	US-08-479-723A-54	Sequence 54, Appl
c	22	14	48.3	44	5	PCT-US94-04310-54	Sequence 54, Appl
	23	14	48.3	47	4	US-09-641-638-1059	Sequence 1059, Ap
c	24	13.8	47.6	17	4	US-09-827-998-283	Sequence 283, App
c	25	13.8	47.6	25	4	US-09-827-998-1105	Sequence 1105, Ap
c	26	13.8	47.6	25	4	US-09-827-998-1106	Sequence 1106, Ap
c	27	13.8	47.6	36	3	US-09-440-001-3	Sequence 3, Appli
c	28	13.8	47.6	36	4	US-09-605-685-3	Sequence 3, Appli
	29	13.8	47.6	47	4	US-09-422-978-639	Sequence 639, App
c	30	13.6	46.9	42	4	US-09-468-872-11	Sequence 11, Appl
	31	13.6	46.9	44	2	US-08-343-443B-39	Sequence 39, Appl
c	32	13.6	46.9	47	4	US-09-422-978-2286	Sequence 2286, Ap
c	33	13.6	46.9	55	4	US-08-956-171E-5024	Sequence 5024, Ap
	34	13.4	46.2	26	3	US-09-247-190-37	Sequence 37, Appl
c	35	13.4	46.2	26	4	US-10-061-658-4	Sequence 4, Appli
c	36	13.4	46.2	35	4	US-09-598-747-32	Sequence 32, Appl
	37	13.4	46.2	41	1	US-08-468-220-28	Sequence 28, Appl
	38	13.4	46.2	41	2	US-08-468-698-28	Sequence 28, Appl
	39	13.4	46.2	41	3	US-08-194-664A-28	Sequence 28, Appl
	40	13.4	46.2	41	5	PCT-US94-01553A-28	Sequence 28, Appl
	41	13.4	46.2	41	5	PCT-US95-10426-28	Sequence 28, Appl
c	42	13.4	46.2	51	1	US-08-328-152A-11	Sequence 11, Appl
	43	13.4	46.2	52	4	US-09-310-463-6	Sequence 6, Appli
	44	13.4	46.2	52	4	US-08-842-248A-6	Sequence 6, Appli
c	45	13.4	46.2	60	3	US-08-478-097A-32	Sequence 32, Appl

ALIGNMENTS

```
RESULT 1
US-08-943-731-336/c
; Sequence 336, Application US/08943731
; Patent No. 6265157
;  GENERAL INFORMATION:
;    APPLICANT:  PROCKOP, DARWIN J.
;    APPLICANT:  SPOTILA, LORETTA D.
;    APPLICANT:  DELTAS, CONSTANTINOS D.
;    APPLICANT:  SEREDA, LARISA
;    APPLICANT:  LARSON, ANDREA W.
;    APPLICANT:  PACK, MICHAEL
;    APPLICANT:  COLIGE, ALAIN
;    APPLICANT:  EARLY, JAMES
;    APPLICANT:  KORKKO, JARMO
;    APPLICANT:  ALA-KOKKO, LEENA, et al.
;    TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR DETECTING
;    TITLE OF INVENTION:  ALTERED TYPE I OR TYPE IX COLLAGEN GENE SEQUENCES
;    NUMBER OF SEQUENCES:  666
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PANITCH SCHWARZE JACOBS & NADEL, P.C.
;      STREET:  ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND
;      STREET:  FLR.
;      CITY:  PHILADELPHIA
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103-7086
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/943,731
;      FILING DATE:  03-OCT-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/212,322
;      FILING DATE:  14-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/803,628
;      FILING DATE:  03-DEC-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DOYLE LEARY Ph.D., KATHRYN
;      REGISTRATION NUMBER:  36,317
;      REFERENCE/DOCKET NUMBER:  9598-27
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-965-1284
;      TELEFAX:  215-567-2991
;      TELEX:  831-494
;  INFORMATION FOR SEQ ID NO:  336:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  25 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-943-731-336

  Query Match            52.4%;  Score 15.2;  DB 3;  Length 25;
  Best Local Similarity  45.0%;  Pred. No. 3e+02;
  Matches    9;  Conservative    8;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          5 AUUCUUUUUGUAAGCCCUAG 24
              |::|: :::|: ||||| |
Db         22 ATTCTCTTTGTGAGCCCTTG 3


RESULT 2
US-08-667-079B-5/c
; Sequence 5, Application US/08667079B
; Patent No. 5789171
;  GENERAL INFORMATION:
;    APPLICANT:  Mark S. Smeltzer
;    TITLE OF INVENTION:  Use of cna, fnbA, fnbB, and hlb Gene Probes for the Strain-Sp
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Benjamin Aaron Adler, MCGREGOR & ADLER, P.C.
;      STREET:  8011 Candle Lane
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Apple Macintosh
;      OPERATING SYSTEM:  Macintosh
;      SOFTWARE:  Microsoft Word for Macintosh
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/667,079B
;      FILING DATE:  June 20, 1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Adler, Benjamin Aaron
;      REGISTRATION NUMBER:  35,423
;      REFERENCE/DOCKET NUMBER:  D5886
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  713-777-2321
;      TELEFAX:  713-777-6908
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  33
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:
;      DESCRIPTION:  other nucleic acid
;    HYPOTHETICAL:  No
;    ANTI-SENSE:  No
;    ORIGINAL SOURCE:
;      STRAIN:
;      INDIVIDUAL ISOLATE:
;      DEVELOPMENTAL STAGE:
;      TISSUE TYPE:
;      CELL TYPE:
;      CELL LINE:
US-08-667-079B-5

  Query Match            52.4%;  Score 15.2;  DB 1;  Length 33;
  Best Local Similarity  39.3%;  Pred. No. 3.1e+02;
  Matches   11;  Conservative    9;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 AUGAUUCUUUUUGUAAGCCCUAGGGGCU 29
              |:||:: :::: |:||   |   ||||:
Db         32 ATGATTGTTTTAGTAATTTCCCCGGGCT 5


RESULT 3
5422260-12/c
;Patent No. 5422260
;    APPLICANT: KAUFMAN, RANDAL J.;PITTMAN, DEBRA D.;TOOLE,JOHN J.
;    TITLE OF INVENTION: HUMAN FACTOR VIII: C MUTEINS
;    NUMBER OF SEQUENCES: 15
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/883,936
;      FILING DATE: 15-MAY-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 279,485
;      FILING DATE: 02-DEC-1988; 09-DEC-1986
;      APPLICATION NUMBER: 939,658
;      FILING DATE: 09-DEC-1986
;      APPLICATION NUMBER: 932,767
;      FILING DATE: 18-NOV-1986
;      APPLICATION NUMBER: 868,410
;      FILING DATE: 29-MAY-1986
;SEQ ID NO:12:
;      LENGTH: 35
5422260-12

  Query Match            51.0%;  Score 14.8;  DB 6;  Length 35;
  Best Local Similarity  42.3%;  Pred. No. 4.9e+02;
  Matches   11;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          4 GAUUCUUUUUGUAAGCCCUAGGGGCU 29
              | ::  ::::| ||||  : |||||:
Db         35 GTTTTCTTTTGAAAGCTTTTGGGGCT 10


RESULT 4
US-09-440-001-1/c
; Sequence 1, Application US/09440001
; Patent No. 6174696
; GENERAL INFORMATION:
;  APPLICANT:  Seman, Leo J.
;  TITLE OF INVENTION:  A METHOD FOR THE DETERMINATION OF HOMOCYSTEINE
;  FILE REFERENCE:  09/440,001
;  CURRENT APPLICATION NUMBER: US/09/440,001
;  CURRENT FILING DATE:  1999-11-12
;  PRIOR APPLICATION NUMBER:  60/108,099
;  PRIOR FILING DATE:  1998-11-12
;  NUMBER OF SEQ ID NOS:  6
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 36
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:
;   OTHER INFORMATION: Oligonucleotide primer
US-09-440-001-1

  Query Match            51.0%;  Score 14.8;  DB 3;  Length 36;
  Best Local Similarity  38.5%;  Pred. No. 4.9e+02;
  Matches   10;  Conservative    9;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 UAUGAUUCUUUUUGUAAGCCCUAGGG 26
              :|: |  |:::::|:  ||| :| ||
Db         33 TATCAAGCTTTTTGTCCGCCATATGG 8


RESULT 5
US-09-605-685-1/c
; Sequence 1, Application US/09605685
; Patent No. 6436658
; GENERAL INFORMATION:
;  APPLICANT: Seman, Leo J.
;  TITLE OF INVENTION: A METHOD FOR THE DETERMINATION OF HOMOCYSTEINE
```

```
;  FILE REFERENCE: 09/440,001
;  CURRENT APPLICATION NUMBER: US/09/605,685
;  CURRENT FILING DATE:  2000-06-26
;  PRIOR APPLICATION NUMBER: 60/108,099
;  PRIOR FILING DATE: 1998-11-12
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 36
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:
;   OTHER INFORMATION: Oligonucleotide primer
US-09-605-685-1

  Query Match             51.0%;  Score 14.8;  DB 4;  Length 36;
  Best Local Similarity   38.5%;  Pred. No. 4.9e+02;
  Matches   10;  Conservative    9;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 UAUGAUUCUUUUUGUAAGCCCUAGGG 26
              :|: |  |:::::|:  ||| :| ||
Db         33 TATCAAGCTTTTTGTCCGCCATATGG 8


RESULT 6
US-09-827-998-1098/c
; Sequence 1098, Application US/09827998
; Patent No. 6656700
; GENERAL INFORMATION:
;  APPLICANT: Gu, Yizhong
;  APPLICANT:  Shannon, Mark
;  TITLE OF INVENTION: NOVEL ISOFORMS OF HUMAN PREGNANCY-ASSOCIATED PROTEIN E
;  FILE REFERENCE: MDhMORF-8
;  CURRENT APPLICATION NUMBER: US/09/827,998
;  CURRENT FILING DATE:  2001-04-06
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  NUMBER OF SEQ ID NOS: 1881
;  SOFTWARE: Aeomica Sequence Listing Engine
; Patent No. 6656700
; SEQ ID NO 1098
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-827-998-1098

  Query Match             49.0%;  Score 14.2;  DB 4;  Length 25;
  Best Local Similarity   42.1%;  Pred. No. 8.7e+02;
  Matches    8;  Conservative    8;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 UCUUUUUGUAAGCCCUAGG 25
              :|:::::|:|  |||:| |
Db         25 TCTTTTTGTAGTCCCTAAG 7


RESULT 7
US-09-827-998-1099/c
; Sequence 1099, Application US/09827998
; Patent No. 6656700
; GENERAL INFORMATION:
;  APPLICANT: Gu, Yizhong
;  APPLICANT:  Shannon, Mark
;  TITLE OF INVENTION: NOVEL ISOFORMS OF HUMAN PREGNANCY-ASSOCIATED PROTEIN E
;  FILE REFERENCE: MDhMORF-8
;  CURRENT APPLICATION NUMBER: US/09/827,998
;  CURRENT FILING DATE:  2001-04-06
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  NUMBER OF SEQ ID NOS: 1881
;  SOFTWARE: Aeomica Sequence Listing Engine
; Patent No. 6656700
; SEQ ID NO 1099
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-827-998-1099

  Query Match             49.0%;  Score 14.2;  DB 4;  Length 25;
  Best Local Similarity   42.1%;  Pred. No. 8.7e+02;
  Matches    8;  Conservative    8;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 UCUUUUUGUAAGCCCUAGG 25
              :|:::::|:|  |||:| |
Db         24 TCTTTTTGTAGTCCCTAAG 6


RESULT 8
US-09-827-998-1100/c
; Sequence 1100, Application US/09827998
; Patent No. 6656700
; GENERAL INFORMATION:
;  APPLICANT: Gu, Yizhong
;  APPLICANT:  Shannon, Mark
;  TITLE OF INVENTION: NOVEL ISOFORMS OF HUMAN PREGNANCY-ASSOCIATED PROTEIN E
;  FILE REFERENCE: MDhMORF-8
;  CURRENT APPLICATION NUMBER: US/09/827,998
;  CURRENT FILING DATE:  2001-04-06
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  NUMBER OF SEQ ID NOS: 1881
;  SOFTWARE: Aeomica Sequence Listing Engine
; Patent No. 6656700
; SEQ ID NO 1100
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-827-998-1100

  Query Match             49.0%;  Score 14.2;  DB 4;  Length 25;
  Best Local Similarity   42.1%;  Pred. No. 8.7e+02;
  Matches    8;  Conservative    8;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 UCUUUUUGUAAGCCCUAGG 25
              :|:::::|:|  |||:| |
Db         23 TCTTTTTGTAGTCCCTAAG 5


RESULT 9
US-09-827-998-1101/c
; Sequence 1101, Application US/09827998
; Patent No. 6656700
; GENERAL INFORMATION:
;  APPLICANT: Gu, Yizhong
;  APPLICANT:  Shannon, Mark
;  TITLE OF INVENTION: NOVEL ISOFORMS OF HUMAN PREGNANCY-ASSOCIATED PROTEIN E
;  FILE REFERENCE: MDhMORF-8
;  CURRENT APPLICATION NUMBER: US/09/827,998
;  CURRENT FILING DATE:  2001-04-06
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  NUMBER OF SEQ ID NOS: 1881
;  SOFTWARE: Aeomica Sequence Listing Engine
; Patent No. 6656700
; SEQ ID NO 1101
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

US-09-827-998-1101

Query Match 49.0%; Score 14.2; DB 4; Length 25;
Best Local Similarity 42.1%; Pred. No. 8.7e+02;
Matches 8; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

```
Qy          7 UCUUUUUGUAAGCCCUAGG 25
              :|:::::|:|  |||:| |
Db         22 TCTTTTTGTAGTCCCTAAG 4
```

RESULT 10
US-09-827-998-1102/c
; Sequence 1102, Application US/09827998
; Patent No. 6656700
; GENERAL INFORMATION:
; APPLICANT: Gu, Yizhong
; APPLICANT: Shannon, Mark
; TITLE OF INVENTION: NOVEL ISOFORMS OF HUMAN PREGNANCY-ASSOCIATED PROTEIN E
; FILE REFERENCE: MDhMORF-8
; CURRENT APPLICATION NUMBER: US/09/827,998
; CURRENT FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 1881
; SOFTWARE: Aeomica Sequence Listing Engine
; Patent No. 6656700
; SEQ ID NO 1102
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-827-998-1102

Query Match 49.0%; Score 14.2; DB 4; Length 25;
Best Local Similarity 42.1%; Pred. No. 8.7e+02;
Matches 8; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

```
Qy          7 UCUUUUUGUAAGCCCUAGG 25
              :|:::::|:|  |||:| |
Db         21 TCTTTTTGTAGTCCCTAAG 3
```

RESULT 11
US-09-827-998-1103/c
; Sequence 1103, Application US/09827998
; Patent No. 6656700
; GENERAL INFORMATION:
; APPLICANT: Gu, Yizhong
; APPLICANT: Shannon, Mark
; TITLE OF INVENTION: NOVEL ISOFORMS OF HUMAN PREGNANCY-ASSOCIATED PROTEIN E
; FILE REFERENCE: MDhMORF-8
; CURRENT APPLICATION NUMBER: US/09/827,998
; CURRENT FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 1881
; SOFTWARE: Aeomica Sequence Listing Engine
; Patent No. 6656700
; SEQ ID NO 1103
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-827-998-1103

Query Match 49.0%; Score 14.2; DB 4; Length 25;
Best Local Similarity 42.1%; Pred. No. 8.7e+02;
Matches 8; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

```
Qy          7 UCUUUUUGUAAGCCCUAGG 25
              :|:::::|:|  |||:| |
Db         20 TCTTTTTGTAGTCCCTAAG 2
```

RESULT 12
US-09-827-998-1104/c
; Sequence 1104, Application US/09827998
; Patent No. 6656700
; GENERAL INFORMATION:
; APPLICANT: Gu, Yizhong
; APPLICANT: Shannon, Mark
; TITLE OF INVENTION: NOVEL ISOFORMS OF HUMAN PREGNANCY-ASSOCIATED PROTEIN E
; FILE REFERENCE: MDhMORF-8
; CURRENT APPLICATION NUMBER: US/09/827,998
; CURRENT FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 1881
; SOFTWARE: Aeomica Sequence Listing Engine
; Patent No. 6656700
; SEQ ID NO 1104
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-827-998-1104

Query Match 49.0%; Score 14.2; DB 4; Length 25;
Best Local Similarity 42.1%; Pred. No. 8.7e+02;
Matches 8; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

```
Qy          7 UCUUUUUGUAAGCCCUAGG 25
              :|:::::|:|  |||:| |
Db         19 TCTTTTTGTAGTCCCTAAG 1
```

RESULT 13
US-09-690-146A-5
; Sequence 5, Application US/09690146A
; Patent No. 6485937
; GENERAL INFORMATION:
; APPLICANT: Palhan, Vikas
; APPLICANT: Roeder, Robert
; TITLE OF INVENTION: System for Rapid Generation of Recombinant
; TITLE OF INVENTION: Baculovirus-Based Expression Vectors for Silkworm Larvae
; FILE REFERENCE: 7529/1G164-US1
; CURRENT APPLICATION NUMBER: US/09/690,146A
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/159,707
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: VP28 Reverse Primer
US-09-690-146A-5

Query Match 48.3%; Score 14; DB 4; Length 30;
Best Local Similarity 45.5%; Pred. No. 1.1e+03;
Matches 10; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

```
Qy          5 AUUCUUUUUGUAAGCCCUAGGG 26
              |::   :::|:|| || :||||
Db          2 ATTAAATTTGTAATCCTTAGGG 23
```

RESULT 14
US-09-690-146A-7/c
; Sequence 7, Application US/09690146A

```
; Patent No. 6485937
; GENERAL INFORMATION:
;  APPLICANT: Palhan, Vikas
;  APPLICANT:  Roeder, Robert
;  TITLE OF INVENTION: System for Rapid Generation of Recombinant
;  TITLE OF INVENTION:  Baculovirus-Based Expression Vectors for Silkworm Larvae
;  FILE REFERENCE: 7529/1G164-US1
;  CURRENT APPLICATION NUMBER: US/09/690,146A
;  CURRENT FILING DATE:  2001-06-01
;  PRIOR APPLICATION NUMBER: 60/159,707
;  PRIOR FILING DATE: 1999-10-15
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
;   LENGTH: 30
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthesized oligonucleotide
US-09-690-146A-7

  Query Match             48.3%;  Score 14;  DB 4;  Length 30;
  Best Local Similarity   45.5%;  Pred. No. 1.1e+03;
  Matches   10;  Conservative    7;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          5 AUUCUUUUUGUAAGCCCUAGGG 26
              |::   :::|:|| || :||||
Db         29 ATTAAATTTGTAATCCTTAGGG 8


RESULT 15
US-08-049-264C-55
; Sequence 55, Application US/08049264C
; Patent No. 5518901
;  GENERAL INFORMATION:
;    APPLICANT:  Murtagh, James J.
;    TITLE OF INVENTION:  METHODS FOR NUCLEIC ACID DETECTION,
;    TITLE OF INVENTION:  SEQUENCING AND CLONING USING EXONUCLEASE
;    NUMBER OF SEQUENCES:  75
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NEEDLE & ROSENBERG, P.C.
;      STREET:  Suite 1200, The Candler Bldg., 127
;      STREET:  Peachtree Street N.E.
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/049,264C
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Perryman, David G.
;      REGISTRATION NUMBER:  33,438
;      REFERENCE/DOCKET NUMBER:  1313.001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404) 688-0770
;      TELEFAX:  (404) 688-9880
;  INFORMATION FOR SEQ ID NO:  55:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  37 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-049-264C-55

  Query Match             48.3%;  Score 14;  DB 1;  Length 37;
  Best Local Similarity   40.9%;  Pred. No. 1.2e+03;
  Matches    9;  Conservative    8;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          6 UUCUUUUUGUAAGCCCUAGGGG 27
              :: ::::: :|| |||  ||||
Db          8 TTTTTTTTTAAACCCGGGGGG 29


Search completed: March 23, 2004, 17:20:34
Job time : 58.3333 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 14:06:04 ; Search time 237.667 Seconds
(without alignments)
451.369 Million cell updates/sec

Title: US-09-310-844C-24
Perfect score: 29
Sequence: 1 uaugauucuuuuuguaagcccuaggggcu 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2438257 seqs, 1849576744 residues

Total number of hits satisfying chosen parameters: 1514106

Minimum DB seq length: 0
Maximum DB seq length: 70

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		16	55.2	41	14	US-10-005-956-275	Sequence 275, App
c 2		15.6	53.8	38	14	US-10-122-706-20	Sequence 20, Appl
3		15.6	53.8	59	14	US-10-122-706-19	Sequence 19, Appl
4		15.4	53.1	53	9	US-09-983-965-4754	Sequence 4754, Ap
5		15	51.7	25	14	US-10-098-263B-37315	Sequence 37315, A
c 6		14.8	51.0	43	14	US-10-032-585-694	Sequence 694, App
7		14.8	51.0	50	15	US-10-131-827-2022	Sequence 2022, Ap
c 8		14.6	50.3	65	14	US-10-032-585-2227	Sequence 2227, Ap
9		14.4	49.7	41	14	US-10-005-956-212	Sequence 212, App
c 10		14.2	49.0	25	9	US-09-827-998-1098	Sequence 1098, Ap
c 11		14.2	49.0	25	9	US-09-827-998-1099	Sequence 1099, Ap
c 12		14.2	49.0	25	9	US-09-827-998-1100	Sequence 1100, Ap
c 13		14.2	49.0	25	9	US-09-827-998-1101	Sequence 1101, Ap
c 14		14.2	49.0	25	9	US-09-827-998-1102	Sequence 1102, Ap
c 15		14.2	49.0	25	9	US-09-827-998-1103	Sequence 1103, Ap
c 16		14.2	49.0	25	9	US-09-827-998-1104	Sequence 1104, Ap
17		14.2	49.0	50	15	US-10-402-365-49	Sequence 49, Appl
c 18		14.2	49.0	56	10	US-09-800-130A-8	Sequence 8, Appli
c 19		14.2	49.0	56	14	US-10-413-909-8	Sequence 8, Appli
20		14.2	49.0	60	10	US-09-908-975-13840	Sequence 13840, A
21		14	48.3	47	15	US-10-170-097-1059	Sequence 1059, Ap
c 22		14	48.3	60	10	US-09-908-975-6479	Sequence 6479, Ap
c 23		14	48.3	60	10	US-09-908-975-12187	Sequence 12187, A
c 24		13.8	47.6	17	9	US-09-827-998-283	Sequence 283, App
c 25		13.8	47.6	25	9	US-09-827-998-1105	Sequence 1105, Ap
c 26		13.8	47.6	25	9	US-09-827-998-1106	Sequence 1106, Ap
c 27		13.8	47.6	25	9	US-09-872-462-287	Sequence 287, App
28		13.8	47.6	25	14	US-10-215-112-14195	Sequence 14195, A
c 29		13.8	47.6	25	14	US-10-098-263B-8809	Sequence 8809, Ap
c 30		13.8	47.6	25	14	US-10-098-263B-82359	Sequence 82359, A
31		13.8	47.6	30	9	US-09-465-802-6	Sequence 6, Appli
32		13.8	47.6	47	15	US-10-349-143-639	Sequence 639, App
c 33		13.8	47.6	60	10	US-09-908-975-6521	Sequence 6521, Ap
34		13.8	47.6	60	10	US-09-908-975-22550	Sequence 22550, A
c 35		13.8	47.6	65	14	US-10-032-585-158	Sequence 158, App
36		13.6	46.9	31	9	US-09-817-360-11	Sequence 11, Appl
c 37		13.6	46.9	31	9	US-09-801-274-1228	Sequence 1228, Ap
c 38		13.6	46.9	31	9	US-09-745-008-7	Sequence 7, Appli
c 39		13.6	46.9	31	10	US-09-848-754A-7495	Sequence 7495, Ap
c 40		13.6	46.9	31	10	US-09-780-164-2256	Sequence 2256, Ap
c 41		13.6	46.9	31	10	US-09-740-332-6085	Sequence 6085, Ap
c 42		13.6	46.9	31	10	US-09-817-879-6085	Sequence 6085, Ap
c 43		13.6	46.9	47	15	US-10-349-143-2286	Sequence 2286, Ap
c 44		13.6	46.9	50	15	US-10-131-827-4574	Sequence 4574, Ap
c 45		13.6	46.9	50	15	US-10-131-827-7620	Sequence 7620, Ap

ALIGNMENTS

RESULT 1
US-10-005-956-275
; Sequence 275, Application US/10005956
; Publication No. US20030113726A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
; FILE REFERENCE: D0053NP
; CURRENT APPLICATION NUMBER: US/10/005,956
; CURRENT FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: 60/251,015
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 60/263,678
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/273,037
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 1579
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 275
; LENGTH: 41
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-005-956-275

Query Match 55.2%; Score 16; DB 14; Length 41;
Best Local Similarity 50.0%; Pred. No. 1.2e+03;
Matches 12; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

```
Qy          3 UGAUUCUUUUUGUAAGCCCUAGGG 26
              :||  |:::::| ||| ||:|| |
Db          2 TGACCCTTTTTGCAAGTCCTAGTG 25
```

RESULT 2
US-10-122-706-20/c
; Sequence 20, Application US/10122706
; Publication No. US20030119012A1
; GENERAL INFORMATION:

```
;  APPLICANT: Srinivasan, Maithreyan
;  APPLICANT:  Reifler, Michael
;  TITLE OF INVENTION: Sulfurylase-Luciferase Fusion Proteins
;  FILE REFERENCE: 21465-504
;  CURRENT APPLICATION NUMBER: US/10/122,706
;  CURRENT FILING DATE:  2002-07-01
;  PRIOR APPLICATION NUMBER: 60/335,949
;  PRIOR FILING DATE: 2001-10-30
;  NUMBER OF SEQ ID NOS: 31
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
;   LENGTH: 38
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: primer
US-10-122-706-20

  Query Match             53.8%;  Score 15.6;  DB 14;  Length 38;
  Best Local Similarity   45.5%;  Pred. No. 1.9e+03;
  Matches   10;  Conservative    8;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          7 UCUUUUUGUAAGCCCUAGGGGC 28
              : :::::|: |||| :|| |||
Db         25 TGTTTTTGTCAGCCATAGCGGC 4


RESULT 3
US-10-122-706-19
; Sequence 19, Application US/10122706
; Publication No. US20030119012A1
; GENERAL INFORMATION:
;  APPLICANT: Srinivasan, Maithreyan
;  APPLICANT:  Reifler, Michael
;  TITLE OF INVENTION: Sulfurylase-Luciferase Fusion Proteins
;  FILE REFERENCE: 21465-504
;  CURRENT APPLICATION NUMBER: US/10/122,706
;  CURRENT FILING DATE:  2002-07-01
;  PRIOR APPLICATION NUMBER: 60/335,949
;  PRIOR FILING DATE: 2001-10-30
;  NUMBER OF SEQ ID NOS: 31
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
;   LENGTH: 59
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: primer
US-10-122-706-19

  Query Match             53.8%;  Score 15.6;  DB 14;  Length 59;
  Best Local Similarity   45.5%;  Pred. No. 2.1e+03;
  Matches   10;  Conservative    8;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          7 UCUUUUUGUAAGCCCUAGGGGC 28
              : :::::|: |||| :|| |||
Db         14 TGTTTTTGTCAGCCATAGCGGC 35


RESULT 4
US-09-983-965-4754
; Sequence 4754, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 37-21(10297)C
;  CURRENT APPLICATION NUMBER: US/09/983,965
;  CURRENT FILING DATE:  2001-10-26
;  PRIOR APPLICATION NUMBER: US 09/465,231
;  PRIOR FILING DATE: 1999-12-15
;  PRIOR APPLICATION NUMBER: US 60/113,678
;  PRIOR FILING DATE: 1998-12-17
;  NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 4754
;   LENGTH: 53
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   FEATURE:
;   OTHER INFORMATION: Clone ID: 18-LIB34-011-Q1-E1-E5
US-09-983-965-4754

  Query Match             53.1%;  Score 15.4;  DB 9;  Length 53;
  Best Local Similarity   40.0%;  Pred. No. 2.5e+03;
  Matches   10;  Conservative    9;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          5 AUUCUUUUUGUAAGCCCUAGGGGCU 29
              |::|::: :|:  ||| :  ||||:
Db         28 ATTCTTTGTGTTTGCCTTCAGGGCT 52


RESULT 5
US-10-098-263B-37315
; Sequence 37315, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
;  APPLICANT: Mittman, Michael
;  TITLE OF INVENTION: Human Microarray
;  FILE REFERENCE: 3118.1
;  CURRENT APPLICATION NUMBER: US/10/098,263B
;  CURRENT FILING DATE:  2003-01-08
;  PRIOR APPLICATION NUMBER: 60/276,759
;  PRIOR FILING DATE: 2001-03-16
;  NUMBER OF SEQ ID NOS: 131066
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 37315
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-098-263B-37315

  Query Match             51.7%;  Score 15;  DB 14;  Length 25;
  Best Local Similarity   47.8%;  Pred. No. 3.2e+03;
  Matches   11;  Conservative    7;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          3 UGAUUCUUUUUGUAAGCCCUAGG 25
              : |  | ::::|:| ||||:|||
Db          3 TCACACATTTTGTACGCCCTAGG 25


RESULT 6
US-10-032-585-694/c
; Sequence 694, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
;  APPLICANT: Terry, Roemer D.
;  APPLICANT:  Bo, Jiang
;  APPLICANT:  Charles, Boone
;  APPLICANT:  Howard, Bussey
;  TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
;  FILE REFERENCE: 10182-005-999
;  CURRENT APPLICATION NUMBER: US/10/032,585
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 8000
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 694
;   LENGTH: 43
;   TYPE: DNA
;   ORGANISM: Candida albicans
US-10-032-585-694

  Query Match             51.0%;  Score 14.8;  DB 14;  Length 43;
```

```
  Best Local Similarity   38.9%;  Pred. No. 4.5e+03;
  Matches    7;  Conservative    9;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 UAUGAUUCUUUUUGUAAG 18
              :|:|| :|:::::|: ||
Db         23 TATGAATCTTTTTGTTAG 6


RESULT 7
US-10-131-827-2022
; Sequence 2022, Application US/10131827
; Publication No. US20040009479A1
; GENERAL INFORMATION:
;  APPLICANT: Wohlgemuth, Jay
;  APPLICANT:  Fry, Kirk
;  APPLICANT:  Woodward, Robert
;  APPLICANT:  Ly, Ngoc
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE
;  TITLE OF INVENTION:  CHRONIC INFLAMMATORY DISEASES
;  FILE REFERENCE: 506612000120
;  CURRENT APPLICATION NUMBER: US/10/131,827
;  CURRENT FILING DATE:  2002-09-06
;  PRIOR APPLICATION NUMBER: US 10/006,290
;  PRIOR FILING DATE: 2001-10-22
;  PRIOR APPLICATION NUMBER: US 60/296,764
;  PRIOR FILING DATE: 2001-06-08
;  NUMBER OF SEQ ID NOS: 9090
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2022
;   LENGTH: 50
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-131-827-2022

  Query Match             51.0%;  Score 14.8;  DB 15;  Length 50;
  Best Local Similarity   38.9%;  Pred. No. 4.7e+03;
  Matches    7;  Conservative    9;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 AUGAUUCUUUUUGUAAGC 19
              |:||:: ::::: :||||
Db          3 ATGATTATTTTTCTAAGC 20


RESULT 8
US-10-032-585-2227/c
; Sequence 2227, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
;  APPLICANT: Terry, Roemer D.
;  APPLICANT:  Bo, Jiang
;  APPLICANT:  Charles, Boone
;  APPLICANT:  Howard, Bussey
;  TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
;  FILE REFERENCE: 10182-005-999
;  CURRENT APPLICATION NUMBER: US/10/032,585
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 8000
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2227
;   LENGTH: 65
;   TYPE: DNA
;   ORGANISM: Candida albicans
US-10-032-585-2227

  Query Match             50.3%;  Score 14.6;  DB 14;  Length 65;
  Best Local Similarity   27.6%;  Pred. No. 6.2e+03;
  Matches    8;  Conservative   12;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 UAUGAUUCUUUUUGUAAGCCCUAGGGGCU 29
              :|:  :: :::::|:||   |:||   |:
Db         57 TATTTTTTTTTTTGTAAAGACTAGAACCT 29


RESULT 9
US-10-005-956-212
; Sequence 212, Application US/10005956
; Publication No. US20030113726A1
; GENERAL INFORMATION:
;  APPLICANT: Bristol-Myers Squibb Company
;  TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
;  FILE REFERENCE: D0053NP
;  CURRENT APPLICATION NUMBER: US/10/005,956
;  CURRENT FILING DATE:  2001-12-03
;  PRIOR APPLICATION NUMBER: 60/251,015
;  PRIOR FILING DATE: 2000-12-04
;  PRIOR APPLICATION NUMBER: 60/263,678
;  PRIOR FILING DATE: 2001-01-23
;  PRIOR APPLICATION NUMBER: 60/273,037
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 1579
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 212
;   LENGTH: 41
;   TYPE: DNA
;   ORGANISM: homo sapiens
US-10-005-956-212

  Query Match             49.7%;  Score 14.4;  DB 14;  Length 41;
  Best Local Similarity   50.0%;  Pred. No. 6.8e+03;
  Matches   12;  Conservative    6;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          3 UGAUUCUUUUUGUAAGCCCUAGGG 26
              :||  |:::::| ||| || || |
Db          2 TGACCCTTTTTGCAAGTCCCAGTG 25


RESULT 10
US-09-827-998-1098/c
; Sequence 1098, Application US/09827998
; Patent No. US20020102252A1
; GENERAL INFORMATION:
;  APPLICANT: Gu, Yizhong
;  APPLICANT:  Shannon, Mark
;  TITLE OF INVENTION: NOVEL ISOFORMS OF HUMAN PREGNANCY-ASSOCIATED PROTEIN E
;  FILE REFERENCE: MDhMORF-8
;  CURRENT APPLICATION NUMBER: US/09/827,998
;  CURRENT FILING DATE:  2001-04-06
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  NUMBER OF SEQ ID NOS: 1881
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 1098
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-827-998-1098

  Query Match             49.0%;  Score 14.2;  DB 9;  Length 25;
  Best Local Similarity   42.1%;  Pred. No. 7.5e+03;
  Matches    8;  Conservative    8;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 UCUUUUUGUAAGCCCUAGG 25
              :|:::::|:|  |||:| |
Db         25 TCTTTTTGTAGTCCCTAAG 7


RESULT 11
US-09-827-998-1099/c
; Sequence 1099, Application US/09827998
; Patent No. US20020102252A1
; GENERAL INFORMATION:
;  APPLICANT: Gu, Yizhong
;  APPLICANT:  Shannon, Mark
;  TITLE OF INVENTION: NOVEL ISOFORMS OF HUMAN PREGNANCY-ASSOCIATED PROTEIN E
```

```
;  FILE REFERENCE: MDhMORF-8
;  CURRENT APPLICATION NUMBER: US/09/827,998
;  CURRENT FILING DATE:  2001-04-06
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  NUMBER OF SEQ ID NOS: 1881
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 1099
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-827-998-1099

  Query Match             49.0%;  Score 14.2;  DB 9;  Length 25;
  Best Local Similarity   42.1%;  Pred. No. 7.5e+03;
  Matches    8;  Conservative    8;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 UCUUUUUGUAAGCCCUAGG 25
              :|:::::|:|  |||:| |
Db         24 TCTTTTTGTAGTCCCTAAG 6


RESULT 12
US-09-827-998-1100/c
; Sequence 1100, Application US/09827998
; Patent No. US20020102252A1
; GENERAL INFORMATION:
;  APPLICANT: Gu, Yizhong
;  APPLICANT:  Shannon, Mark
;  TITLE OF INVENTION: NOVEL ISOFORMS OF HUMAN PREGNANCY-ASSOCIATED PROTEIN E
;  FILE REFERENCE: MDhMORF-8
;  CURRENT APPLICATION NUMBER: US/09/827,998
;  CURRENT FILING DATE:  2001-04-06
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  NUMBER OF SEQ ID NOS: 1881
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 1100
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-827-998-1100

  Query Match             49.0%;  Score 14.2;  DB 9;  Length 25;
  Best Local Similarity   42.1%;  Pred. No. 7.5e+03;
  Matches    8;  Conservative    8;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 UCUUUUUGUAAGCCCUAGG 25
              :|:::::|:|  |||:| |
Db         23 TCTTTTTGTAGTCCCTAAG 5


RESULT 13
US-09-827-998-1101/c
; Sequence 1101, Application US/09827998
; Patent No. US20020102252A1
; GENERAL INFORMATION:
;  APPLICANT: Gu, Yizhong
;  APPLICANT:  Shannon, Mark
;  TITLE OF INVENTION: NOVEL ISOFORMS OF HUMAN PREGNANCY-ASSOCIATED PROTEIN E
;  FILE REFERENCE: MDhMORF-8
;  CURRENT APPLICATION NUMBER: US/09/827,998
;  CURRENT FILING DATE:  2001-04-06
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  NUMBER OF SEQ ID NOS: 1881
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 1101
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-827-998-1101

  Query Match             49.0%;  Score 14.2;  DB 9;  Length 25;
  Best Local Similarity   42.1%;  Pred. No. 7.5e+03;
  Matches    8;  Conservative    8;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 UCUUUUUGUAAGCCCUAGG 25
              :|:::::|:|  |||:| |
Db         22 TCTTTTTGTAGTCCCTAAG 4


RESULT 14
US-09-827-998-1102/c
; Sequence 1102, Application US/09827998
; Patent No. US20020102252A1
; GENERAL INFORMATION:
;  APPLICANT: Gu, Yizhong
;  APPLICANT:  Shannon, Mark
;  TITLE OF INVENTION: NOVEL ISOFORMS OF HUMAN PREGNANCY-ASSOCIATED PROTEIN E
;  FILE REFERENCE: MDhMORF-8
;  CURRENT APPLICATION NUMBER: US/09/827,998
;  CURRENT FILING DATE:  2001-04-06
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  NUMBER OF SEQ ID NOS: 1881
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 1102
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-827-998-1102

  Query Match             49.0%;  Score 14.2;  DB 9;  Length 25;
  Best Local Similarity   42.1%;  Pred. No. 7.5e+03;
  Matches    8;  Conservative    8;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 UCUUUUUGUAAGCCCUAGG 25
              :|:::::|:|  |||:| |
Db         21 TCTTTTTGTAGTCCCTAAG 3


RESULT 15
US-09-827-998-1103/c
; Sequence 1103, Application US/09827998
; Patent No. US20020102252A1
; GENERAL INFORMATION:
;  APPLICANT: Gu, Yizhong
;  APPLICANT:  Shannon, Mark
;  TITLE OF INVENTION: NOVEL ISOFORMS OF HUMAN PREGNANCY-ASSOCIATED PROTEIN E
;  FILE REFERENCE: MDhMORF-8
;  CURRENT APPLICATION NUMBER: US/09/827,998
;  CURRENT FILING DATE:  2001-04-06
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  NUMBER OF SEQ ID NOS: 1881
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 1103
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-827-998-1103

  Query Match             49.0%;  Score 14.2;  DB 9;  Length 25;
  Best Local Similarity   42.1%;  Pred. No. 7.5e+03;
  Matches    8;  Conservative    8;  Mismatches    3;  Indels    0;  Gaps    0;
```

```
Qy          7 UCUUUUUGUAAGCCCUAGG 25
              :|::::::|:|   |||:| |
Db         20 TCTTTTTGTAGTCCCTAAG 2
```

Search completed: March 23, 2004, 17:17:34
Job time : 242 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 14:02:34 ; Search time 1997.33 Seconds
(without alignments)
433.580 Million cell updates/sec

Title: US-09-310-844C-24
Perfect score: 29
Sequence: 1 uaugauucuuuuuguaagcccuaggggcu 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 289680

Minimum DB seq length: 0
Maximum DB seq length: 70

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	16.2	55.9	57	9	AI561770	AI561770 vv65b08.x
c	2	15.4	53.1	61	29	CG590967	CG590967 OST245023
c	3	15.4	53.1	66	29	TA123H02P	AL463084 T. brucei
c	4	15.2	52.4	41	14	CB210991	CB210991 OML01271
	5	15.2	52.4	63	29	BX533881	BX533881 Arabidops
c	6	15.2	52.4	69	29	CG732697	CG732697 1119150C1
c	7	14.8	51.0	25	28	AZ993079	AZ993079 2M0277P20
	8	14.8	51.0	60	10	BE871815	BE871815 601447803
c	9	14.8	51.0	64	29	BX003595	BX003595 Arabidops
c	10	14.8	51.0	66	14	CD925111	CD925111 G750.115O
	11	14.8	51.0	69	12	BM128463	BM128463 if15c05.x
	12	14.6	50.3	49	29	BX287070	BX287070 Arabidops
	13	14.6	50.3	52	12	BG236504	BG236504 nai44c06.
c	14	14.6	50.3	62	28	BH911891	BH911891 SALK_0727
c	15	14.6	50.3	63	29	CG563472	CG563472 OST186777
c	16	14.6	50.3	64	9	AI139668	AI139668 qc28h07.x
	17	14.6	50.3	64	28	AZ808107	AZ808107 2M0071O24
	18	14.6	50.3	66	28	AZ440181	AZ440181 1M0231E10
c	19	14.4	49.7	31	28	BH910631	BH910631 SALK_0607
c	20	14.4	49.7	43	28	AZ597048	AZ597048 1M0410K10
	21	14.4	49.7	44	29	AL771575	AL771575 Arabidops
	22	14.4	49.7	54	12	BI865449	BI865449 ft22h02.x
	23	14.4	49.7	60	9	AL595218	AL595218 AL595218
c	24	14.4	49.7	63	12	BG362434	BG362434 gb72b09.y
c	25	14.4	49.7	65	29	CG664319	CG664319 OST451176
c	26	14.4	49.7	67	9	AI584052	AI584052 ts13b02.x
c	27	14.4	49.7	67	10	BE027305	BE027305 EtESTea89
	28	14.4	49.7	67	28	BZ289657	BZ289657 SALK_0230
c	29	14.4	49.7	68	29	CG581901	CG581901 OST222599
c	30	14.4	49.7	69	10	BE647308	BE647308 UI-M-BH1-
	31	14.2	49.0	52	10	AW686481	AW686481 NF038E05N
c	32	14.2	49.0	59	28	B00509	B00509 cSRL-115b2-
	33	14.2	49.0	61	9	AI318033	AI318033 ta75g02.x
	34	14.2	49.0	67	29	BX004510	BX004510 Arabidops
	35	14.2	49.0	70	13	BU063954	BU063954 Fgr_3_M18
c	36	14	48.3	52	10	BF637245	BF637245 NF002F08L
c	37	14	48.3	65	12	BI094834	BI094834 EST-CD34N
c	38	14	48.3	67	9	AA936041	AA936041 nz53f10.s
	39	14	48.3	67	28	BH855810	BH855810 SALK_0845
	40	13.8	47.6	37	29	AL951243	AL951243 Arabidops
	41	13.8	47.6	39	28	BH909815	BH909815 SALK_0561
	42	13.8	47.6	40	28	BH857340	BH857340 SALK_0764
	43	13.8	47.6	40	28	BH857342	BH857342 SALK_0764
	44	13.8	47.6	40	29	CG779591	CG779591 1123034H1
	45	13.8	47.6	43	28	AZ484548	AZ484548 1M0311N02

ALIGNMENTS

RESULT 1
AI561770
LOCUS AI561770 57 bp mRNA linear EST 25-MAR-1999
DEFINITION vv65b08.x1 Stratagene mouse skin (#937313) Mus musculus cDNA clone IMAGE:1227255 3', mRNA sequence.
ACCESSION AI561770
VERSION AI561770.1 GI:4513115
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 57)
AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE The WashU-NCI Mouse EST Project 1999
JOURNAL Unpublished (1999)
COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the

```
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:652847
            This clone was previously sequenced on the 5' end only, this new
            data is from the 3' end
            High quality sequence stop: 51.
FEATURES             Location/Qualifiers
     source          1. .57
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1227255"
                     /sex="females"
                     /tissue_type="whole skin"
                     /dev_stage="11 weeks old"
                     /lab_host="SOLR (kanamycin resistant)"
                     /clone_lib="Stratagene mouse skin (#937313)"
                     /note="Organ: skin; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. Whole skin from 11 week old C57BL/6 female mice.
                     Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5'
                     adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor
                     sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"
ORIGIN

  Query Match             55.9%;  Score 16.2;  DB 9;  Length 57;
  Best Local Similarity   37.9%;  Pred. No. 7.6e+04;
  Matches   11;  Conservative   10;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 UAUGAUUCUUUUUGUAAGCCCUAGGGGCU 29
              : :|| :| :::: :|| || : |||| :
Db         25 TTTGAATCCTTTTCTAATCCATGGGGGGT 53


RESULT 2
CG590967/c
LOCUS       CG590967                  61 bp    DNA     linear   GSS 02-OCT-2003
DEFINITION  OST245023 Mus musculus 129Sv/Ev Mus musculus genomic clone
            OST245023, genomic survey sequence.
ACCESSION   CG590967
VERSION     CG590967.1  GI:37398378
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 61)
  AUTHORS   Zambrowicz,B.P., Abuin,A., Ramirez-Solis,R., Richter,L.J.,
            Piggott,J., BeltrandelRio,H., Buxton,E.C., Edwards,J., Finch,R.A.,
            Friddle,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaing,C.,
            Key,B.W. Jr., Kipp,P., Kohlhauff,B., Ma,Z.-Q., Markesich,D.,
            Payne,R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,Z.-Z.,
            Sparks,M.J., Van Sligtenhorst,I., Vogel,P., Walke,W., Xu,N.,
            Zhu,Q., Person,C. and Sands,A.T.
  TITLE     Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap
            screen to identify potential targets for therapeutic intervention
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
COMMENT     Contact: Zambrowicz BP
            OmniBank
            Lexicon Genetics Incorporated
            4000 Research Forest Drive, The Woodlands, TX 77381, USA
            Email: materials@lexgen.com
            Gene trap sequence tag generated by 3' RACE from mouse ES cells as
            described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)
            Class: Gene Trap.
FEATURES             Location/Qualifiers
     source          1. .61
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="129SV/EV"
                     /db_xref="taxon:10090"
                     /clone="OST245023"
                     /cell_type="embryonic stem cell"
                     /clone_lib="Mus musculus 129Sv/Ev"
ORIGIN

  Query Match             53.1%;  Score 15.4;  DB 29;  Length 61;
  Best Local Similarity   50.0%;  Pred. No. 1.5e+05;
  Matches   13;  Conservative    6;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          3 UGAUUCUUUUUGUAAGCCCUAGGGGC 28
              :||  |: :::|  |||||:|| | |
Db         58 TGAGCCTNTTTGCCAGCCCTAGTGCC 33


RESULT 3
TA123H02P/c
LOCUS       TA123H02P                 66 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  T. brucei sheared genomic DNA clone 123h02, forward sequence,
            genomic survey sequence.
ACCESSION   AL463084
VERSION     AL463084.1  GI:11833690
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 66)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nh1@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb).  The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects.  In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES             Location/Qualifiers
     source          1. .66
                     /organism="Trypanosoma brucei"
                     /mol_type="genomic DNA"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
                     /clone="123h02"
ORIGIN

  Query Match             53.1%;  Score 15.4;  DB 29;  Length 66;
  Best Local Similarity   36.0%;  Pred. No. 1.5e+05;
  Matches    9;  Conservative   10;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 UAUGAUUCUUUUUGUAAGCCCUAGG 25
              :|:||:: :::::    | |||:| |
Db         30 TATGATTTTTTTTCAGAACCCTAAG 6


RESULT 4
CB210991/c
LOCUS       CB210991                  41 bp    mRNA    linear   EST 05-FEB-2003
DEFINITION  OML01271 Oryza minuta HybriZAP-2.1 XR library Oryza minuta cDNA 5',
            mRNA sequence.
ACCESSION   CB210991
VERSION     CB210991.1  GI:28257082
KEYWORDS    EST.
SOURCE      Oryza minuta
  ORGANISM  Oryza minuta
```

```
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 41)
  AUTHORS   Shin,J.S.
  TITLE     Oryza minuta HybriZAP-2.1 XR library
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Jeong Sheop Shin
            Plant Molecular Genetics
            Graduate School of Biotechnology, University of Korea
            136-701 Anam-dong 5/1 Seoul, Korea
            Tel: 00 82 2 3290 3430
            Fax: 00 82 2 927 9028
            Email: jsshin@kuccnx.korea.ac.kr.
FEATURES             Location/Qualifiers
     source          1. .41
                     /organism="Oryza minuta"
                     /mol_type="mRNA"
                     /db_xref="taxon:63629"
                     /dev_stage="4-weeks after germination"
                     /clone_lib="Oryza minuta HybriZAP-2.1 XR library"
                     /note="Organ: immature leaf"
ORIGIN

  Query Match             52.4%;  Score 15.2;  DB 14;  Length 41;
  Best Local Similarity   39.3%;  Pred. No. 1.9e+05;
  Matches   11;  Conservative    9;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 UAUGAUUCUUUUUGUAAGCCCUAGGGGC 28
              :|:| ::|:: :  : |||||: | | |
Db         28 TATGCTTCTTGTGCTTAGCCCTCGTGCC 1


RESULT 5
BX533881
LOCUS       BX533881                 63 bp    DNA     linear   GSS 03-JUN-2003
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence GK-504G04-019705,
            genomic survey sequence.
ACCESSION   BX533881
VERSION     BX533881.1  GI:31411011
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Strizhov,N., Li,Y., Rosso,M., Viehoever,P., Dekker,K., Saedler,H.
            and Weisshaar,B.
  TITLE     A pipeline for automated high-throughput generation of FSTs
            (flanking sequence tags) from Arabidopsis thaliana T-DNA
            transformed lines
  JOURNAL   Unpublished
REFERENCE   2
  AUTHORS   Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B.
  TITLE     A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)
            for flanking sequence tag based reverse genetics
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 63)
  AUTHORS   Strizhov,N., Li,Y., Rosso,M. and Weisshaar,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-2003) Weisshaar B., Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT     This sequence is recovered from the left border of the T-DNA. It
            indicates an insertion close to or within gene At5g40230. The
            sequences are generated at the MPI for Plant Breeding Research in
            the context of the GABI-Kat project. GABI-Kat is part of the German
            Plant Genomics program designated 'GABI'. Information on line
            availability can be found at:
            http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES             Location/Qualifiers
     source          1. .63
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="GK-504G04-019705"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                     /note="PCR was performed on DNA from Arabidopsis thaliana
                     plants (T1) which were transformed with the T-DNA from
                     vector pAC161. The lines contain one or more T-DNA
                     insertions. The DNA fragment(s) resulting from the PCR
                     were directly sequenced to determine the genomic sequence
                     flanking the insertion. Sequences displaying significant
                     similarity to the A. thalaina nuclear genome sequence were
                     processed for submission. T-DNA derived sequences were
                     removed"
ORIGIN

  Query Match             52.4%;  Score 15.2;  DB 29;  Length 63;
  Best Local Similarity   39.3%;  Pred. No. 1.8e+05;
  Matches   11;  Conservative    9;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 AUGAUUCUUUUUGUAAGCCCUAGGGGCU 29
              |: | : :::::  || |||:| ||| :
Db         21 ATAAATATTTTTTCAAACCCTATGGGAT 48


RESULT 6
CG732697/c
LOCUS       CG732697                 69 bp    DNA     linear   GSS 20-OCT-2003
DEFINITION  1119150C11.y1 1119 - RescueMu Grid AA Zea mays genomic, genomic
            survey sequence.
ACCESSION   CG732697
VERSION     CG732697.1  GI:37775189
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 69)
  AUTHORS   Walbot,V.
  TITLE     Maize genomic sequences found using engineered RescueMu transposon
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 1119150  row: C  column: 11
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .69
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="mixed background W23/A188/B73/K55"
                     /db_xref="taxon:4577"
                     /tissue_type="leaf"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="1119 - RescueMu Grid AA"
                     /note="Organ: leaf; Vector: RescueMu (engineered from
                     pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
                     RescueMu is a 4.9 kb, modified maize Mu transposon
                     designed to allow plasmid rescue from total genomic DNA.
                     Mu elements insert preferentially into transcription
                     units. For more information on RescueMu, go to the web
                     site 'www.zmdb.iastate.edu' and follow the links for
                     'RescueMu.' Grid AA was grown at UC San Diego in 2002. DNA
                     was extracted from leaf strips, double digested using
                     BamHI and BglII, and ligated to form circular plasmids.
                     DH10B cells were transformed and then screened on LB
                     plates with ampicillin."
```

```
ORIGIN

  Query Match             52.4%;  Score 15.2;  DB 29;  Length 69;
  Best Local Similarity   35.7%;  Pred. No. 1.7e+05;
  Matches   10;  Conservative   10;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 UAUGAUUCUUUUUGUAAGCCCUAGGGGC 28
              :|: |:: : :::|:|   |   : |||||
Db         39 TATTATTATATTTGTATTCTGTTGGGGC 12


RESULT 7
AZ993079/c
LOCUS       AZ993079                  25 bp    DNA     linear   GSS 27-APR-2001
DEFINITION  2M0277P20R Mouse 10kb plasmid UUGC2M library Mus musculus genomic
            clone UUGC2M0277P20 R, genomic survey sequence.
ACCESSION   AZ993079
VERSION     AZ993079.1  GI:13864306
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 25)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0277  row: P  column: 20
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 25.
FEATURES             Location/Qualifiers
     source          1. .25
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0277P20"
                     /sex="Female"
                     /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC2M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (female) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."

ORIGIN

  Query Match             51.0%;  Score 14.8;  DB 28;  Length 25;
  Best Local Similarity   72.2%;  Pred. No. 2.9e+05;
  Matches   13;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         11 UUUGUAAGCCCUAGGGGC 28
              :::| |||||| ||||||
Db         21 TTTGCAAGCCCAAGGGGC 4


RESULT 8
BE871815
LOCUS       BE871815                  60 bp    mRNA    linear   EST 20-OCT-2000
DEFINITION  601447803F1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3851880 5',
            mRNA sequence.
ACCESSION   BE871815
VERSION     BE871815.1  GI:10320591
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9573  row: e  column: 01
            High quality sequence stop: 60.
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3851880"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_65"
                     /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.8 kb.   Library constructed by Life
                     Technologies.  "
ORIGIN

  Query Match             51.0%;  Score 14.8;  DB 10;  Length 60;
  Best Local Similarity   38.9%;  Pred. No. 2.5e+05;
  Matches    7;  Conservative    9;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 AUGAUUCUUUUUGUAAGC 19
              |:||:: ::::: :||||
Db         10 ATGATTATTTTTCTAAGC 27


RESULT 9
BX003595/c
LOCUS       BX003595                  64 bp    DNA     linear   GSS 04-DEC-2002
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence GK-373C10-017165,
            genomic survey sequence.
ACCESSION   BX003595
VERSION     BX003595.1  GI:26188555
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

```
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Strizhov,N., Li,Y., Rosso,M., Viehoever,P., Dekker,K., Saedler,H.
            and Weisshaar,B.
  TITLE     A pipeline for automated high-throughput generation of FSTs
            (flanking sequence tags) from Arabidopsis thaliana T-DNA
            transformed lines
  JOURNAL   Unpublished
REFERENCE   2
  AUTHORS   Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B.
  TITLE     A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)
            for flanking sequence tag based reverse genetics
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 64)
  AUTHORS   Strizhov,N., Li,Y., Rosso,M. and Weisshaar,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-DEC-2002) Weisshaar B., Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT     This sequence is recovered from the left border of the T-DNA. It
            indicates an insertion close to or within gene At1g33610. The
            sequences are generated at the MPI for Plant Breeding Research in
            the context of the GABI-Kat project. GABI-Kat is part of the German
            Plant Genomics program designated 'GABI'. Information on line
            availability can be found at:
            http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES             Location/Qualifiers
     source          1. .64
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="GK-373C10-017165"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                     /note="PCR was performed on DNA from Arabidopsis thaliana
                     plants (T1) which were transformed with the T-DNA from
                     vector pAC161. The lines contain one or more T-DNA
                     insertions. The DNA fragment(s) resulting from the PCR
                     were directly sequenced to determine the genomic sequence
                     flanking the insertion. Sequences displaying significant
                     similarity to the A. thalaina nuclear genome sequence were
                     processed for submission. T-DNA derived sequences were
                     removed"
ORIGIN

  Query Match             51.0%;  Score 14.8;  DB 29;  Length 64;
  Best Local Similarity   34.6%;  Pred. No. 2.5e+05;
  Matches    9;  Conservative    10;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 AUGAUUCUUUUUGUAAGCCCUAGGGG 27
              |:  :: ::::: :| ||| : ||||
Db         60 ATCTTTTTTTTTATATGCCTTTGGGG 35


RESULT 10
CD925111/c
LOCUS       CD925111                   66 bp    mRNA    linear   EST 15-JUL-2003
DEFINITION  G750.115O07F010706 G750 Triticum aestivum cDNA clone G750115O07,
            mRNA sequence.
ACCESSION   CD925111
VERSION     CD925111.1  GI:32772875
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 66)
  AUTHORS   Genoplante.
  TITLE     Genoplante, a major partnership french program in plant genomics
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Genoplante
            Genoplante
            93, rue Henri Rochefort 91025 EVRY CEDEX France
            Tel: 33 1 69 47 54 00
            Fax: 33 1 69 47 54 10
            This sequence has been generated in the framework of the french
            plant genomics programme 'Genoplante' (http://www.genoplante.com
            and http://genoplante-info.infobiogen.fr).
FEATURES             Location/Qualifiers
     source          1. .66
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="recital"
                     /db_xref="taxon:4565"
                     /clone="G750115O07"
                     /tissue_type="grain (750 degrees per day after
                     pollination)"
                     /clone_lib="G750"
ORIGIN

  Query Match             51.0%;  Score 14.8;  DB 14;  Length 66;
  Best Local Similarity   38.5%;  Pred. No. 2.5e+05;
  Matches   10;  Conservative    9;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 AUGAUUCUUUUUGUAAGCCCUAGGGG 27
              |:   : ::::: :||| ||:||| |
Db         45 ATAGATGTTTTTTTAAGTCCTAGGCG 20


RESULT 11
BM128463
LOCUS       BM128463                   69 bp    mRNA    linear   EST 12-MAR-2002
DEFINITION  if15c05.x1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
            cDNA clone IMAGE:5676297 3', mRNA sequence.
ACCESSION   BM128463
VERSION     BM128463.1  GI:17123015
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 69)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Douglas Melton DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Juliana Brown
            (brown@fas.harvard.edu) This sequence now available from the IMAGE
            consortium, for clone orders contact: info@image.llnl.gov.
FEATURES             Location/Qualifiers
     source          1. .69
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5676297"
                     /sex="Both"
                     /tissue_type="Islets of Langerhans"
                     /dev_stage="Adult"
                     /lab_host="DH10B"
                     /clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
                     /note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1;
```

Site_2: Sal 1; Starting library constructed using SuperScript Plasmid Library kit (Life Technologies). cDNA made by oligo-dT priming. Size-selected by column fractionation; average insert size 1.08 kb. Library was amplified once on solid support and plasmid DNA from library was prepared. The library DNA was normalized by method #4 from Bonaldo, Lennon, and Soares 1996 Genome Research 6:791-806; 0.5 microgram single-stranded library plasmid DNA was mixed with 5 micrograms PCR product representing library inserts and hybridized to an Ecot of 20. Single-stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."

ORIGIN

```
  Query Match             51.0%;  Score 14.8;  DB 12;  Length 69;
  Best Local Similarity   30.8%;  Pred. No. 2.5e+05;
  Matches    8;  Conservative   11;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 UAUGAUUCUUUUUGUAAGCCCUAGGG 26
              : :  :: ::::: :  ||||:||||
Db          8 TTTTTTTTTTTTTCTGGGCCCTAGGG 33
```

RESULT 12
BX287070
LOCUS BX287070 49 bp DNA linear GSS 07-MAR-2003
DEFINITION Arabidopsis thaliana T-DNA flanking sequence GK-396F11-018295, genomic survey sequence.
ACCESSION BX287070
VERSION BX287070.1 GI:28886066
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Strizhov,N., Li,Y., Rosso,M., Viehoever,P., Dekker,K., Saedler,H. and Weisshaar,B.
TITLE A pipeline for automated high-throughput generation of FSTs (flanking sequence tags) from Arabidopsis thaliana T-DNA transformed lines
JOURNAL Unpublished
REFERENCE 2
AUTHORS Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B.
TITLE A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat) for flanking sequence tag based reverse genetics
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 49)
AUTHORS Strizhov,N., Li,Y., Rosso,M. and Weisshaar,B.
TITLE Direct Submission
JOURNAL Submitted (07-MAR-2003) Weisshaar B., Max-Planck-Institut fuer Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT This sequence is recovered from the left border of the T-DNA. It indicates an insertion within the locus defined by clone MZN24. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German Plant Genomics program designated 'GABI'. Information on line availability can be found at:
http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES Location/Qualifiers
source 1. .49
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="GK-396F11-018295"
/clone_lib="Arabidopsis thaliana T-DNA insertion lines"
/note="PCR was performed on DNA from Arabidopsis thaliana plants (T1) which were transformed with the T-DNA from vector pAC161. The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. Sequences displaying significant similarity to the A. thalaina nuclear genome sequence were processed for submission. T-DNA derived sequences were removed"

ORIGIN

```
  Query Match             50.3%;  Score 14.6;  DB 29;  Length 49;
  Best Local Similarity   41.4%;  Pred. No. 3.1e+05;
  Matches   12;  Conservative    8;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 UAUGAUUCUUUUUGUAAGCCCUAGGGGCU 29
              :| | ::   : :|:||| ||:| ||| :
Db          8 TAAGCTTAAATGTGTAAGGCCTATGGGAT 36
```

RESULT 13
BG236504
LOCUS BG236504 52 bp mRNA linear EST 12-FEB-2001
DEFINITION nai44c06.x1 NCI_CGAP_HN20 Homo sapiens cDNA clone IMAGE:4262795 3', mRNA sequence.
ACCESSION BG236504
VERSION BG236504.1 GI:12750351
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 52)
AUTHORS NCI/NIDR-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute / National Institute of Dental Research, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -40UP from Gibco.
FEATURES Location/Qualifiers
source 1. .52
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4262795"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_HN20"
/note="Organ: normal head/neck tissue; Vector: pAMP1; mRNA made from head/neck tissue, cDNA made by oligo-dT priming. Directionally cloned into UDG sites. Size-selected on agarose gel, average insert size 300 bp. Primary library. cDNA Library Preparation: David B. Krizman, Ph.D."

ORIGIN

```
  Query Match             50.3%;  Score 14.6;  DB 12;  Length 52;
  Best Local Similarity   38.1%;  Pred. No. 3.1e+05;
  Matches    8;  Conservative    9;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          6 UUCUUUUUGUAAGCCCUAGGG 26
              :: ::::: :|||  |:||||
Db          1 TTTTTTTTTAAGTTCTAGGG 21
```

RESULT 14
BH911891/c
LOCUS BH911891 62 bp DNA linear GSS 04-SEP-2002
DEFINITION SALK_072729.50.65.x Arabidopsis thaliana TDNA insertion lines Arabidopsis thaliana genomic clone SALK_072729.50.65.x, genomic survey sequence.

```
ACCESSION   BH911891
VERSION     BH911891.1  GI:22724824
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 62)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
            Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
            Shinn,P., Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA. This sequence lies within 300 bases of the 5' end of
            At4g13720.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .62
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_072729.50.65.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
ORIGIN

  Query Match             50.3%;  Score 14.6;  DB 28;  Length 62;
  Best Local Similarity   31.0%;  Pred. No. 3e+05;
  Matches    9;  Conservative   11;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 UAUGAUUCUUUUUGUAAGCCCUAGGGGCU 29
              :|: |:: ::::  :| || |:||   |:
Db         46 TATAATTTTTTTACTAGGCGCTAGACCCT 18


RESULT 15
CG563472/c
LOCUS       CG563472                  63 bp    DNA     linear   GSS 01-OCT-2003
DEFINITION  OST186777 Mus musculus 129Sv/Ev Mus musculus genomic clone
            OST186777, genomic survey sequence.
ACCESSION   CG563472
VERSION     CG563472.1  GI:37350059
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 63)
  AUTHORS   Zambrowicz,B.P., Abuin,A., Ramirez-Solis,R., Richter,L.J.,
            Piggott,J., BeltrandelRio,H., Buxton,E.C., Edwards,J., Finch,R.A.,
            Friddle,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaing,C.,
            Key,B.W. Jr., Kipp,P., Kohlhauff,B., Ma,Z.-Q., Markesich,D.,
            Payne,R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,Z.-Z.,
            Sparks,M.J., Van Sligtenhorst,I., Vogel,P., Walke,W., Xu,N.,
            Zhu,Q., Person,C. and Sands,A.T.
  TITLE     Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap
            screen to identify potential targets for therapeutic intervention
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
COMMENT     Contact: Zambrowicz BP
            OmniBank
            Lexicon Genetics Incorporated
            4000 Research Forest Drive, The Woodlands, TX 77381, USA
            Email: materials@lexgen.com
            Gene trap sequence tag generated by 3' RACE from mouse ES cells as
            described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)
            Class: Gene Trap.
FEATURES             Location/Qualifiers
     source          1. .63
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="129SV/EV"
                     /db_xref="taxon:10090"
                     /clone="OST186777"
                     /cell_type="embryonic stem cell"
                     /clone_lib="Mus musculus 129Sv/Ev"
ORIGIN

  Query Match             50.3%;  Score 14.6;  DB 29;  Length 63;
  Best Local Similarity   47.6%;  Pred. No. 3e+05;
  Matches   10;  Conservative    7;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 GAUUCUUUUUGUAAGCCCUAG 24
              | ::|: : :|:||||||: |
Db         45 GTTTCTCTGTGTAAGCCCTGG 25


Search completed: March 23, 2004, 17:05:52
Job time : 2014.33 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 11:44:09 ; Search time 633.333 Seconds
(without alignments)
1984.655 Million cell updates/sec

Title: US-09-310-844C-25
Perfect score: 29
Sequence: 1 aaagauucuuuuuguaagccccaagggcu 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 1733942

Minimum DB seq length: 0
Maximum DB seq length: 70

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		29	100.0	42	6	BD274272	BD274272 Identific
2		29	100.0	42	6	BD274273	BD274273 Identific
3		29	100.0	42	6	BD274280	BD274280 Identific
4		29	100.0	42	6	BD274281	BD274281 Identific
5		28	96.6	46	6	BD274241	BD274241 Identific
6		28	96.6	46	6	BD274242	BD274242 Identific
7		28	96.6	46	6	BD274243	BD274243 Identific
8		28	96.6	46	6	BD274258	BD274258 Identific
9		28	96.6	46	6	BD274259	BD274259 Identific
10		28	96.6	46	6	BD274260	BD274260 Identific
11		24.8	85.5	42	6	BD274270	BD274270 Identific
12		24.8	85.5	42	6	BD274278	BD274278 Identific
13		23.8	82.1	46	6	BD274238	BD274238 Identific
14		23.8	82.1	46	6	BD274256	BD274256 Identific
15		23.2	80.0	42	6	BD274275	BD274275 Identific
16		23.2	80.0	42	6	BD274283	BD274283 Identific
17		22.2	76.6	46	6	BD274240	BD274240 Identific
18		22.2	76.6	46	6	BD274249	BD274249 Identific
19		22.2	76.6	46	6	BD274252	BD274252 Identific
20		22.2	76.6	46	6	BD274253	BD274253 Identific
21		22.2	76.6	46	6	BD274257	BD274257 Identific
22		22.2	76.6	46	6	BD274265	BD274265 Identific
23		22.2	76.6	46	6	BD274268	BD274268 Identific
24		22.2	76.6	46	6	BD274269	BD274269 Identific
25		21.2	73.1	42	6	BD274271	BD274271 Identific
26		21.2	73.1	42	6	BD274279	BD274279 Identific
27		21.2	73.1	46	6	BD274247	BD274247 Identific
28		21.2	73.1	46	6	BD274263	BD274263 Identific
29		20	69.0	46	6	BD274237	BD274237 Identific
30		20	69.0	46	6	BD274251	BD274251 Identific
31		20	69.0	46	6	BD274255	BD274255 Identific
32		20	69.0	46	6	BD274267	BD274267 Identific
33		19.6	67.6	42	6	BD274284	BD274284 Identific
34		18.6	64.1	46	6	BD274246	BD274246 Identific
35		18.6	64.1	46	6	BD274248	BD274248 Identific
36		18.6	64.1	46	6	BD274262	BD274262 Identific
37		18.6	64.1	46	6	BD274264	BD274264 Identific
38		18.4	63.4	42	6	BD274274	BD274274 Identific
39		18.4	63.4	42	6	BD274282	BD274282 Identific
40		18.4	63.4	43	8	ATH553793	AJ553793 Arabidops
41		18	62.1	44	6	BD274277	BD274277 Identific
42	c	16.4	56.6	31	6	AX425989	AX425989 Sequence
43		16.4	56.6	42	6	BD274276	BD274276 Identific
44	c	16.2	55.9	48	6	AX018731	AX018731 Sequence
45	c	15.2	52.4	33	6	AR020509	AR020509 Sequence

ALIGNMENTS

RESULT 1
BD274272
LOCUS BD274272 42 bp DNA linear PAT 17-JUL-2003
DEFINITION Identification of molecular interaction sites in RNA for novel drug discovery.
ACCESSION BD274272
VERSION BD274272.1 GI:33084040
KEYWORDS JP 2002526030-A/239.
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 42)
AUTHORS Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
TITLE Identification of molecular interaction sites in RNA for novel drug discovery

```
  JOURNAL   Patent: JP 2002526030-A 239 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/239
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .42
            FT                   /organism='Artificial Sequence'.
FEATURES              Location/Qualifiers
     source           1. .42
                      /organism="synthetic construct"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:32630"
ORIGIN

  Query Match             100.0%;  Score 29;  DB 6;  Length 42;
  Best Local Similarity   69.0%;  Pred. No. 0.0093;
  Matches   20;  Conservative    9;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
              |||||::|:::::|:|||||||||||||:
Db          4 AAAGATTCTTTTTGTAAGCCCCAAGGGCT 32


RESULT 2
BD274273
LOCUS       BD274273                  42 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274273
VERSION     BD274273.1  GI:33084041
KEYWORDS    JP 2002526030-A/240.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 240 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/240
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .42
            FT                   /organism='Artificial Sequence'.
FEATURES              Location/Qualifiers
     source           1. .42
                      /organism="synthetic construct"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:32630"
ORIGIN

  Query Match             100.0%;  Score 29;  DB 6;  Length 42;
  Best Local Similarity   69.0%;  Pred. No. 0.0093;
  Matches   20;  Conservative    9;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
              |||||::|:::::|:|||||||||||||:
Db          4 AAAGATTCTTTTTGTAAGCCCCAAGGGCT 32


RESULT 3
BD274280
LOCUS       BD274280                  42 bp    RNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274280
VERSION     BD274280.1  GI:33084048
KEYWORDS    JP 2002526030-A/247.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 247 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/247
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .42
            FT                   /organism='Artificial Sequence'.
FEATURES              Location/Qualifiers
     source           1. .42
                      /organism="synthetic construct"
                      /mol_type="genomic RNA"
                      /db_xref="taxon:32630"
ORIGIN

  Query Match             100.0%;  Score 29;  DB 6;  Length 42;
  Best Local Similarity   69.0%;  Pred. No. 0.0093;
  Matches   20;  Conservative    9;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
              |||||::|:::::|:|||||||||||||:
Db          4 AAAGATTCTTTTTGTAAGCCCCAAGGGCT 32


RESULT 4
BD274281
LOCUS       BD274281                  42 bp    RNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274281
VERSION     BD274281.1  GI:33084049
KEYWORDS    JP 2002526030-A/248.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 248 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/248
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .42
            FT                   /organism='Artificial Sequence'.
FEATURES              Location/Qualifiers
```

```
     source          1. .42
                     /organism="synthetic construct"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             100.0%;  Score 29;  DB 6;  Length 42;
  Best Local Similarity   69.0%;  Pred. No. 0.0093;
  Matches   20;  Conservative    9;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
              |||||::|:::::|:|||||||||||||:
Db          4 AAAGATTCTTTTTGTAAGCCCCAAGGGCT 32


RESULT 5
BD274241
LOCUS       BD274241                  46 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274241
VERSION     BD274241.1  GI:33084009
KEYWORDS    JP 2002526030-A/208.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 208 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/208
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .46
            FT                   /organism='Artificial Sequence'.
FEATURES              Location/Qualifiers
     source           1. .46
                      /organism="synthetic construct"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:32630"
ORIGIN

  Query Match             96.6%;  Score 28;  DB 6;  Length 46;
  Best Local Similarity   71.4%;  Pred. No. 0.028;
  Matches   20;  Conservative    8;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 28
              |||||::|:::::|:|||||||||||||
Db         19 AAAGATTCTTTTTGTAAGCCCCAAGGGC 46


RESULT 6
BD274242
LOCUS       BD274242                  46 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274242
VERSION     BD274242.1  GI:33084010
KEYWORDS    JP 2002526030-A/209.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 209 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/209
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .46
            FT                   /organism='Artificial Sequence'.
FEATURES              Location/Qualifiers
     source           1. .46
                      /organism="synthetic construct"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:32630"
ORIGIN

  Query Match             96.6%;  Score 28;  DB 6;  Length 46;
  Best Local Similarity   71.4%;  Pred. No. 0.028;
  Matches   20;  Conservative    8;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 28
              |||||::|:::::|:|||||||||||||
Db         19 AAAGATTCTTTTTGTAAGCCCCAAGGGC 46


RESULT 7
BD274243
LOCUS       BD274243                  46 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274243
VERSION     BD274243.1  GI:33084011
KEYWORDS    JP 2002526030-A/210.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 210 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/210
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .46
            FT                   /organism='Artificial Sequence'.
FEATURES              Location/Qualifiers
     source           1. .46
                      /organism="synthetic construct"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:32630"
ORIGIN

  Query Match             96.6%;  Score 28;  DB 6;  Length 46;
  Best Local Similarity   71.4%;  Pred. No. 0.028;
  Matches   20;  Conservative    8;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 28
              |||||::|:::::|:|||||||||||||
Db         19 AAAGATTCTTTTTGTAAGCCCCAAGGGC 46
```

```
RESULT 8
BD274258
LOCUS       BD274258                  46 bp    RNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274258
VERSION     BD274258.1  GI:33084026
KEYWORDS    JP 2002526030-A/225.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 225 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/225
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .46
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="synthetic construct"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             96.6%;  Score 28;  DB 6;  Length 46;
  Best Local Similarity   71.4%;  Pred. No. 0.028;
  Matches   20;  Conservative    8;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 28
              |||||::|:::::|:||||||||||||||
Db         19 AAAGATTCTTTTTGTAAGCCCCAAGGGC 46


RESULT 9
BD274259
LOCUS       BD274259                  46 bp    RNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274259
VERSION     BD274259.1  GI:33084027
KEYWORDS    JP 2002526030-A/226.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 226 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/226
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .46
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="synthetic construct"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             96.6%;  Score 28;  DB 6;  Length 46;
  Best Local Similarity   71.4%;  Pred. No. 0.028;
  Matches   20;  Conservative    8;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 28
              |||||::|:::::|:||||||||||||||
Db         19 AAAGATTCTTTTTGTAAGCCCCAAGGGC 46


RESULT 10
BD274260
LOCUS       BD274260                  46 bp    RNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274260
VERSION     BD274260.1  GI:33084028
KEYWORDS    JP 2002526030-A/227.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 227 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/227
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .46
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="synthetic construct"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             96.6%;  Score 28;  DB 6;  Length 46;
  Best Local Similarity   71.4%;  Pred. No. 0.028;
  Matches   20;  Conservative    8;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 28
              |||||::|:::::|:||||||||||||||
Db         19 AAAGATTCTTTTTGTAAGCCCCAAGGGC 46


RESULT 11
BD274270
LOCUS       BD274270                  42 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274270
VERSION     BD274270.1  GI:33084038
KEYWORDS    JP 2002526030-A/237.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
```

```
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 237 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/237
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .42
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .42
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             85.5%;  Score 24.8;  DB 6;  Length 42;
  Best Local Similarity   60.7%;  Pred. No. 0.95;
  Matches   17;  Conservative    9;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 AAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
              ||||::|:::::|:|||||| | ||||:
Db          5 AAGATTCTTTTTGTAAGCCCTACGGGCT 32


RESULT 12
BD274278
LOCUS       BD274278                  42 bp    RNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274278
VERSION     BD274278.1  GI:33084046
KEYWORDS    JP 2002526030-A/245.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 245 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/245
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .42
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .42
                     /organism="synthetic construct"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             85.5%;  Score 24.8;  DB 6;  Length 42;
  Best Local Similarity   60.7%;  Pred. No. 0.95;
  Matches   17;  Conservative    9;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 AAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
              ||||::|:::::|:|||||| | ||||:
Db          5 AAGATTCTTTTTGTAAGCCCTACGGGCT 32


RESULT 13
BD274238
LOCUS       BD274238                  46 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274238
VERSION     BD274238.1  GI:33084006
KEYWORDS    JP 2002526030-A/205.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 205 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/205
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .46
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             82.1%;  Score 23.8;  DB 6;  Length 46;
  Best Local Similarity   63.0%;  Pred. No. 2.8;
  Matches   17;  Conservative    8;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 AAGAUUCUUUUUGUAAGCCCCAAGGGC 28
              ||||::|:::::|:|||||| | ||||
Db         20 AAGATTCTTTTTGTAAGCCCTACGGGC 46


RESULT 14
BD274256
LOCUS       BD274256                  46 bp    RNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274256
VERSION     BD274256.1  GI:33084024
KEYWORDS    JP 2002526030-A/223.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 223 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/223
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .46
```

```
           FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="synthetic construct"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match           82.1%;  Score 23.8;  DB 6;  Length 46;
  Best Local Similarity   63.0%;  Pred. No. 2.8;
  Matches   17;  Conservative    8;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 AAGAUUCUUUUUGUAAGCCCCAAGGGC 28
              ||||::|:::::|:|||||| | ||||
Db         20 AAGATTCTTTTTGTAAGCCCTACGGGC 46


RESULT 15
BD274275
LOCUS       BD274275                 42 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Identification of molecular interaction sites in RNA for novel drug
            discovery.
ACCESSION   BD274275
VERSION     BD274275.1  GI:33084043
KEYWORDS    JP 2002526030-A/242.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Ecker,D.J., Sampath,R., Griffey,R. and Mcneil,J.
  TITLE     Identification of molecular interaction sites in RNA for novel drug
            discovery
  JOURNAL   Patent: JP 2002526030-A 242 20-AUG-2002;
            ISIS PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002526030-A/242
            PD   20-AUG-2002
            PF   12-MAY-1999 JP 2000548510
            PR   12-MAY-1998 US    60/085092,12-MAY-1998 US    09/076440 PI
            DAVID J ECKER,RANGA SAMPATH,RICHARD GRIFFEY,JOHN MCNEIL PC
            C12Q1/68,A61K31/7105,A61K48/00,C12N15/09,C12N15/00 CC   Description
            of Artificial Sequence: Novel Sequence FH   Key
            Location/Qualifiers
            FT   source          1. .42
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .42
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match           80.0%;  Score 23.2;  DB 6;  Length 42;
  Best Local Similarity   57.1%;  Pred. No. 5.5;
  Matches   16;  Conservative    9;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 AAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
              | ||::|:::::|:|||||| | ||||:
Db          5 ATGATTCTTTTTGTAAGCCCTAGGGGCT 32


Search completed: March 23, 2004, 15:25:09
Job time : 634.333 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 11:34:35 ; Search time 235.333 Seconds
(without alignments)
523.503 Million cell updates/sec

Title: US-09-310-844C-25
Perfect score: 29
Sequence: 1 aaagauucuuuuuguaagccccaagggcu 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 3369620

Minimum DB seq length: 0
Maximum DB seq length: 70

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	100.0	29	3	AAA70829	Aaa70829 Molecular
2	29	100.0	29	3	AAA70830	Aaa70830 Molecular
3	29	100.0	42	3	AAA71121	Aaa71121 Molecular
4	29	100.0	42	3	AAA71128	Aaa71128 Molecular
5	29	100.0	42	3	AAA71120	Aaa71120 Molecular
6	29	100.0	42	3	AAA71116	Aaa71116 Molecular
7	29	100.0	42	3	AAA71115	Aaa71115 Molecular
8	29	100.0	42	3	AAA71129	Aaa71129 Molecular
9	28	96.6	45	3	AAA70826	Aaa70826 Molecular
10	28	96.6	45	3	AAA70825	Aaa70825 Molecular
11	28	96.6	46	3	AAA71089	Aaa71089 Molecular
12	28	96.6	46	3	AAA71106	Aaa71106 Molecular
13	28	96.6	46	3	AAA71107	Aaa71107 Molecular
14	28	96.6	46	3	AAA71088	Aaa71088 Molecular
15	28	96.6	46	3	AAA71105	Aaa71105 Molecular
16	28	96.6	46	3	AAA71090	Aaa71090 Molecular
17	24.8	85.5	42	3	AAA71113	Aaa71113 Molecular
18	24.8	85.5	42	3	AAA71118	Aaa71118 Molecular
19	24.8	85.5	42	3	AAA71126	Aaa71126 Molecular
20	23.8	82.1	46	3	AAA71085	Aaa71085 Molecular
21	23.8	82.1	46	3	AAA71103	Aaa71103 Molecular
22	23.2	80.0	29	3	AAA70828	Aaa70828 Molecular
23	23.2	80.0	42	3	AAA71123	Aaa71123 Molecular
24	23.2	80.0	42	3	AAA71131	Aaa71131 Molecular
25	22.2	76.6	45	3	AAA70824	Aaa70824 Molecular
26	22.2	76.6	46	3	AAA71087	Aaa71087 Molecular
27	22.2	76.6	46	3	AAA71096	Aaa71096 Molecular
28	22.2	76.6	46	3	AAA71099	Aaa71099 Molecular
29	22.2	76.6	46	3	AAA71100	Aaa71100 Molecular
30	22.2	76.6	46	3	AAA71104	Aaa71104 Molecular
31	21.2	73.1	42	3	AAA71114	Aaa71114 Molecular
32	21.2	73.1	42	3	AAA71119	Aaa71119 Molecular
33	21.2	73.1	42	3	AAA71127	Aaa71127 Molecular
34	21.2	73.1	46	3	AAA71094	Aaa71094 Molecular
35	21.2	73.1	46	3	AAA71110	Aaa71110 Molecular
36	20	69.0	46	3	AAA71098	Aaa71098 Molecular
37	20	69.0	46	3	AAA71102	Aaa71102 Molecular
38	20	69.0	46	3	AAA71084	Aaa71084 Molecular
39	19.6	67.6	42	3	AAA71124	Aaa71124 Molecular
40	19.6	67.6	42	3	AAA71132	Aaa71132 Molecular
41	18.6	64.1	46	3	AAA71111	Aaa71111 Molecular
42	18.6	64.1	46	3	AAA71095	Aaa71095 Molecular
43	18.6	64.1	46	3	AAA71109	Aaa71109 Molecular
44	18.6	64.1	46	3	AAA71093	Aaa71093 Molecular
45	18.4	63.4	42	3	AAA71130	Aaa71130 Molecular

ALIGNMENTS

RESULT 1
AAA70829
ID AAA70829 standard; RNA; 29 BP.
XX
AC AAA70829;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #29.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Mus sp.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Claim 235; Page 235; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The

CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 29 BP; 8 A; 6 C; 6 G; 0 T; 9 U; 0 Other;

Query Match 100.0%; Score 29; DB 3; Length 29;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
||||||||||||||||||||||||||||||
Db 1 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 29

RESULT 2
AAA70830
ID AAA70830 standard; RNA; 29 BP.
XX
AC AAA70830;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #30.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Rattus sp.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Claim 235; Page 235; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 29 BP; 8 A; 6 C; 6 G; 0 T; 9 U; 0 Other;

Query Match 100.0%; Score 29; DB 3; Length 29;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
||||||||||||||||||||||||||||||
Db 1 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 29

RESULT 3
AAA71121
ID AAA71121 standard; DNA; 42 BP.
XX
AC AAA71121;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #127.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 125; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary

CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 42 BP; 13 A; 7 C; 7 G; 15 T; 0 U; 0 Other;

Query Match 100.0%; Score 29; DB 3; Length 42;
Best Local Similarity 69.0%; Pred. No. 0.0028;
Matches 20; Conservative 9; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
||||||::|::::|:||||||||||||:
Db 4 AAAGATTCTTTTTGTAAGCCCCAAGGGCT 32

RESULT 4
AAA71128
ID AAA71128 standard; RNA; 42 BP.
XX
AC AAA71128;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #197.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 126; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 42 BP; 13 A; 7 C; 7 G; 0 T; 15 U; 0 Other;

Query Match 100.0%; Score 29; DB 3; Length 42;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
||||||||||||||||||||||||||||
Db 4 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 32

RESULT 5
AAA71120
ID AAA71120 standard; DNA; 42 BP.
XX
AC AAA71120;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #126.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 125; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an

CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 42 BP; 13 A; 7 C; 7 G; 15 T; 0 U; 0 Other;

Query Match 100.0%; Score 29; DB 3; Length 42;
Best Local Similarity 69.0%; Pred. No. 0.0028;
Matches 20; Conservative 9; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
||||||::|::::|:||||||||||||:
Db 4 AAAGATTCTTTTTGTAAGCCCCAAGGGCT 32

RESULT 6
AAA71116
ID AAA71116 standard; RNA; 42 BP.
XX
AC AAA71116;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #192.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 122; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 42 BP; 13 A; 7 C; 7 G; 0 T; 15 U; 0 Other;

Query Match 100.0%; Score 29; DB 3; Length 42;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
||||||||||||||||||||||||||||||
Db 4 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 32

RESULT 7
AAA71115
ID AAA71115 standard; RNA; 42 BP.
XX
AC AAA71115;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #191.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 122; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4

CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 42 BP; 13 A; 7 C; 7 G; 0 T; 15 U; 0 Other;

Query Match 100.0%; Score 29; DB 3; Length 42;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
||||||||||||||||||||||||||||||
Db 4 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 32

RESULT 8
AAA71129
ID AAA71129 standard; RNA; 42 BP.
XX
AC AAA71129;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #198.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 126; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 42 BP; 13 A; 7 C; 7 G; 0 T; 15 U; 0 Other;

Query Match 100.0%; Score 29; DB 3; Length 42;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
||||||||||||||||||||||||||||||
Db 4 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 32

RESULT 9
AAA70826
ID AAA70826 standard; RNA; 45 BP.
XX
AC AAA70826;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #26.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Rattus sp.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Claim 222; Page 232; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3

CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 45 BP; 14 A; 7 C; 9 G; 0 T; 15 U; 0 Other;

Query Match 96.6%; Score 28; DB 3; Length 45;
Best Local Similarity 100.0%; Pred. No. 0.008;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 28
||||||||||||||||||||||||||||
Db 18 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 45

RESULT 10
AAA70825
ID AAA70825 standard; RNA; 45 BP.
XX
AC AAA70825;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #25.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Mus sp.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Claim 221; Page 232; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 45 BP; 14 A; 7 C; 9 G; 0 T; 15 U; 0 Other;

Query Match 96.6%; Score 28; DB 3; Length 45;
Best Local Similarity 100.0%; Pred. No. 0.008;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 28
||||||||||||||||||||||||||||
Db 18 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 45

RESULT 11
AAA71089
ID AAA71089 standard; DNA; 46 BP.
XX
AC AAA71089;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #112.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 121; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence

CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 46 BP; 14 A; 7 C; 9 G; 16 T; 0 U; 0 Other;

Query Match 96.6%; Score 28; DB 3; Length 46;
Best Local Similarity 71.4%; Pred. No. 0.0081;
Matches 20; Conservative 8; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 28
||||||::|:::::|:||||||||||||||
Db 19 AAAGATTCTTTTTGTAAGCCCCAAGGGC 46

RESULT 12
AAA71106
ID AAA71106 standard; RNA; 46 BP.
XX
AC AAA71106;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #182.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 122; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,
CC particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC or industrial compounds
XX
SQ Sequence 46 BP; 14 A; 7 C; 9 G; 0 T; 16 U; 0 Other;

Query Match 96.6%; Score 28; DB 3; Length 46;
Best Local Similarity 100.0%; Pred. No. 0.0081;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 28
||||||||||||||||||||||||||||
Db 19 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 46

RESULT 13
AAA71107
ID AAA71107 standard; RNA; 46 BP.
XX
AC AAA71107;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #183.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US010361.
XX
PR 12-MAY-1998; 98US-00076404.
PR 12-MAY-1998; 98US-0085092P.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, Mcneil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,
PT agricultural and industrial compounds.
XX
PS Example 7; Fig 122; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses 3-
CC dimensional representations of the biomolecule and a library of compounds
CC and comprises (a) identifying at least one molecular interaction site of
CC the target RNA; (b) generating in silico a virtual library of compounds
CC predicted or calculated to interact with the molecular interaction site;
CC and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC with members of the virtual library of compounds to generate a hierarchy
CC of the compounds ranked in accordance with their respective ability to
CC form physical interactions with the molecular interaction site. The
CC method also describes (1) RNA comprising a joined sequence of at least 24
CC nucleotides but not more than 70 nucleotides and having secondary
CC structure defined by: (a) 3 nucleotides forming a first side of a first
CC double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC internal loop region; (c) 4 nucleotides forming a first side of a second
CC ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC nucleotides forming a second side of the second ds region; (f) 4
CC nucleotides forming a second side of the internal loop region; and (g) 3
CC nucleotides forming a second side of the first ds region; (2) a purified
CC and isolated RNA fragment comprising the human sequence
CC UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC used for identifying agents which modulate the activity of biomolecules,

```
CC   particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC   or industrial compounds
XX
SQ   Sequence 46 BP; 14 A; 7 C; 9 G; 0 T; 16 U; 0 Other;

  Query Match             96.6%;  Score 28;  DB 3;  Length 46;
  Best Local Similarity   100.0%;  Pred. No. 0.0081;
  Matches   28;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 28
              ||||||||||||||||||||||||||||
Db         19 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 46


RESULT 14
AAA71088
ID   AAA71088 standard; DNA; 46 BP.
XX
AC   AAA71088;
XX
DT   27-APR-2001  (first entry)
XX
DE   Molecular interaction site DNA #111.
XX
KW   Modulator; identification; molecular interaction; virtual library; ss.
XX
OS   Unidentified.
XX
PN   WO9958947-A2.
XX
PD   18-NOV-1999.
XX
PF   12-MAY-1999;   99WO-US010361.
XX
PR   12-MAY-1998;   98US-00076404.
PR   12-MAY-1998;   98US-0085092P.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Ecker DJ,  Griffey R,  Crooke ST,  Sampath R,  Swayze E,  Mohan V;
PI   Hofstadler S,  Mcneil J;
XX
DR   WPI; 2000-086439/07.
XX
PT   Identifying compounds which modulate activity of target biomolecules,
PT   used to provide compounds which can be used as pharmacological,
PT   agricultural and industrial compounds.
XX
PS   Example 7; Fig 121; 405pp; English.
XX
CC   This invention describes a novel method for identifying compounds which
CC   modulate the activity of a target biomolecule. The method uses 3-
CC   dimensional representations of the biomolecule and a library of compounds
CC   and comprises (a) identifying at least one molecular interaction site of
CC   the target RNA; (b) generating in silico a virtual library of compounds
CC   predicted or calculated to interact with the molecular interaction site;
CC   and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC   with members of the virtual library of compounds to generate a hierarchy
CC   of the compounds ranked in accordance with their respective ability to
CC   form physical interactions with the molecular interaction site. The
CC   method also describes (1) RNA comprising a joined sequence of at least 24
CC   nucleotides but not more than 70 nucleotides and having secondary
CC   structure defined by: (a) 3 nucleotides forming a first side of a first
CC   double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC   internal loop region; (c) 4 nucleotides forming a first side of a second
CC   ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC   nucleotides forming a second side of the second ds region; (f) 4
CC   nucleotides forming a second side of the internal loop region; and (g) 3
CC   nucleotides forming a second side of the first ds region; (2) a purified
CC   and isolated RNA fragment comprising the human sequence
CC   UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC   used for identifying agents which modulate the activity of biomolecules,
CC   particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC   or industrial compounds
XX
SQ   Sequence 46 BP; 14 A; 7 C; 9 G; 16 T; 0 U; 0 Other;

  Query Match             96.6%;  Score 28;  DB 3;  Length 46;
  Best Local Similarity   71.4%;  Pred. No. 0.0081;
  Matches   20;  Conservative    8;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 28
              |||||::|:::::|:||||||||||||||
Db         19 AAAGATTCTTTTTGTAAGCCCCAAGGGC 46


RESULT 15
AAA71105
ID   AAA71105 standard; RNA; 46 BP.
XX
AC   AAA71105;
XX
DT   27-APR-2001  (first entry)
XX
DE   Molecular interaction site RNA #181.
XX
KW   Modulator; identification; molecular interaction; virtual library; ss.
XX
OS   Unidentified.
XX
PN   WO9958947-A2.
XX
PD   18-NOV-1999.
XX
PF   12-MAY-1999;   99WO-US010361.
XX
PR   12-MAY-1998;   98US-00076404.
PR   12-MAY-1998;   98US-0085092P.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Ecker DJ,  Griffey R,  Crooke ST,  Sampath R,  Swayze E,  Mohan V;
PI   Hofstadler S,  Mcneil J;
XX
DR   WPI; 2000-086439/07.
XX
PT   Identifying compounds which modulate activity of target biomolecules,
PT   used to provide compounds which can be used as pharmacological,
PT   agricultural and industrial compounds.
XX
PS   Example 7; Fig 122; 405pp; English.
XX
CC   This invention describes a novel method for identifying compounds which
CC   modulate the activity of a target biomolecule. The method uses 3-
CC   dimensional representations of the biomolecule and a library of compounds
CC   and comprises (a) identifying at least one molecular interaction site of
CC   the target RNA; (b) generating in silico a virtual library of compounds
CC   predicted or calculated to interact with the molecular interaction site;
CC   and (c) comparing 3-dimensional (3-D) representations of the target RNA
CC   with members of the virtual library of compounds to generate a hierarchy
CC   of the compounds ranked in accordance with their respective ability to
CC   form physical interactions with the molecular interaction site. The
CC   method also describes (1) RNA comprising a joined sequence of at least 24
CC   nucleotides but not more than 70 nucleotides and having secondary
CC   structure defined by: (a) 3 nucleotides forming a first side of a first
CC   double stranded (ds) region; (b) 2 nucleotides forming a first side of an
CC   internal loop region; (c) 4 nucleotides forming a first side of a second
CC   ds region; (d) 4 or 5 nucleotides forming an end loop region; (e) 4
CC   nucleotides forming a second side of the second ds region; (f) 4
CC   nucleotides forming a second side of the internal loop region; and (g) 3
CC   nucleotides forming a second side of the first ds region; (2) a purified
CC   and isolated RNA fragment comprising the human sequence
CC   UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The methods and products can be
CC   used for identifying agents which modulate the activity of biomolecules,
CC   particularly RNA. Such agents can be used as pharmaceutical, agricultural
CC   or industrial compounds
```

XX
SQ Sequence 46 BP; 14 A; 7 C; 9 G; 0 T; 16 U; 0 Other;

Query Match 96.6%; Score 28; DB 3; Length 46;
Best Local Similarity 100.0%; Pred. No. 0.0081;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 28
              ||||||||||||||||||||||||||||
Db         19 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 46
```

Search completed: March 23, 2004, 14:53:14
Job time : 235.333 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 14:41:54 ; Search time 55.3333 Seconds
(without alignments)
290.848 Million cell updates/sec

Title: US-09-310-844C-25
Perfect score: 29
Sequence: 1 aaagauucuuuuuguaagccccaagggcu 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 895828

Minimum DB seq length: 0
Maximum DB seq length: 70

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	15.2	52.4	33	1	US-08-667-079B-5	Sequence 5, Appli
2	15.2	52.4	47	4	US-09-422-978-1097	Sequence 1097, Ap
c 3	14.8	51.0	29	1	US-08-219-633-25	Sequence 25, Appl
c 4	14.8	51.0	29	1	US-08-515-236-25	Sequence 25, Appl
c 5	14.8	51.0	29	1	US-08-761-950-25	Sequence 25, Appl
c 6	14.8	51.0	29	2	US-08-632-575B-39	Sequence 39, Appl
c 7	14.8	51.0	29	3	US-09-327-229-31	Sequence 31, Appl
c 8	14.8	51.0	29	4	US-09-199-542B-39	Sequence 39, Appl
c 9	14.8	51.0	29	5	PCT-US95-12608-31	Sequence 31, Appl
c 10	14.8	51.0	32	2	US-08-632-575B-59	Sequence 59, Appl
c 11	14.8	51.0	32	4	US-09-199-542B-59	Sequence 59, Appl
12	14.6	50.3	25	4	US-09-063-733A-18	Sequence 18, Appl
c 13	14.6	50.3	53	2	US-08-486-969-46	Sequence 46, Appl
c 14	14.2	49.0	25	4	US-09-827-998-1098	Sequence 1098, Ap
c 15	14.2	49.0	25	4	US-09-827-998-1099	Sequence 1099, Ap
c 16	14.2	49.0	25	4	US-09-827-998-1100	Sequence 1100, Ap
c 17	14.2	49.0	25	4	US-09-827-998-1101	Sequence 1101, Ap
c 18	14.2	49.0	25	4	US-09-827-998-1102	Sequence 1102, Ap
c 19	14.2	49.0	25	4	US-09-827-998-1103	Sequence 1103, Ap
c 20	14.2	49.0	25	4	US-09-827-998-1104	Sequence 1104, Ap
21	14.2	49.0	47	4	US-09-422-978-96	Sequence 96, Appl
22	14.2	49.0	69	2	US-08-410-654B-30	Sequence 30, Appl
23	14.2	49.0	69	2	US-08-474-851-30	Sequence 30, Appl
24	14.2	49.0	69	2	US-08-481-560-30	Sequence 30, Appl
25	14	48.3	41	4	US-09-571-774-2	Sequence 2, Appli
26	14	48.3	41	4	US-09-852-385-2	Sequence 2, Appli
c 27	13.8	47.6	25	3	US-08-943-731-336	Sequence 336, App
c 28	13.8	47.6	33	4	US-09-199-542B-76	Sequence 76, Appl
c 29	13.8	47.6	47	4	US-09-671-317-663	Sequence 663, App
c 30	13.8	47.6	50	4	US-09-428-082B-401	Sequence 401, App
31	13.8	47.6	57	4	US-09-428-082B-414	Sequence 414, App
c 32	13.8	47.6	60	4	US-09-428-082B-415	Sequence 415, App
33	13.8	47.6	61	4	US-09-428-082B-400	Sequence 400, App
34	13.6	46.9	41	4	US-09-565-156A-2	Sequence 2, Appli
c 35	13.6	46.9	47	4	US-09-422-978-1843	Sequence 1843, Ap
c 36	13.6	46.9	47	4	US-09-402-266B-10	Sequence 10, Appl
37	13.6	46.9	52	4	US-09-310-463-6	Sequence 6, Appli
38	13.6	46.9	52	4	US-08-842-248A-6	Sequence 6, Appli
39	13.4	46.2	32	3	US-08-718-738-16	Sequence 16, Appl
40	13.4	46.2	32	3	US-09-221-844-16	Sequence 16, Appl
41	13.4	46.2	32	5	PCT-US95-03323A-16	Sequence 16, Appl
c 42	13.4	46.2	40	4	US-09-428-082B-418	Sequence 418, App
c 43	13.4	46.2	46	1	US-08-171-389-42	Sequence 42, Appl
c 44	13.4	46.2	46	1	US-08-171-389-45	Sequence 45, Appl
c 45	13.4	46.2	46	1	US-08-123-936-42	Sequence 42, Appl

ALIGNMENTS

RESULT 1
US-08-667-079B-5/c
; Sequence 5, Application US/08667079B
; Patent No. 5789171
; GENERAL INFORMATION:
; APPLICANT: Mark S. Smeltzer
; TITLE OF INVENTION: Use of cna, fnbA, fnbB, and hlb Gene Probes for the Strain-Sp
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Benjamin Aaron Adler, MCGREGOR & ADLER, P.C.
; STREET: 8011 Candle Lane
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word for Macintosh
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/667,079B
; FILING DATE: June 20, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Adler, Benjamin Aaron
; REGISTRATION NUMBER: 35,423
; REFERENCE/DOCKET NUMBER: D5886
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-777-2321
; TELEFAX: 713-777-6908
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: other nucleic acid
; HYPOTHETICAL: No
; ANTI-SENSE: No
; ORIGINAL SOURCE:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
US-08-667-079B-5

```
  Query Match             52.4%;  Score 15.2;  DB 1;  Length 33;
  Best Local Similarity   42.9%;  Pred. No. 4e+02;
  Matches   12;  Conservative    8;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 AAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
              | ||:: :::: |:||   ||  ||||:
Db         32 ATGATTGTTTTAGTAATTTCCCCGGGCT 5


RESULT 2
US-09-422-978-1097
; Sequence 1097, Application US/09422978
; Patent No. 6537751
; GENERAL INFORMATION:
;  APPLICANT: Cohen, Daniel
;  APPLICANT: Blumenfeld, Marta
;  APPLICANT: Chumakov, Ilya
;  TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
;  FILE REFERENCE: GENSET.020CP1
;  CURRENT APPLICATION NUMBER: US/09/422,978
;  CURRENT FILING DATE: 1999-10-20
;  EARLIER APPLICATION NUMBER: US 09/298,850
;  EARLIER FILING DATE: 1999-04-21
;  EARLIER APPLICATION NUMBER: US 60/109,732
;  EARLIER FILING DATE: 1998-11-23
;  EARLIER APPLICATION NUMBER: US 60/082,614
;  EARLIER FILING DATE: 1998-04-21
;  NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 1097
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 24
;   OTHER INFORMATION: 99-2043-220 : polymorphic base A or T
US-09-422-978-1097

  Query Match             52.4%;  Score 15.2;  DB 4;  Length 47;
  Best Local Similarity   50.0%;  Pred. No. 4.3e+02;
  Matches   11;  Conservative    7;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          3 AGAUUCUUUUUGUAAGCCCCAA 24
              ||| :|:::::|:|| | || |
Db         15 AGACTCTTTWTGTAACCTCCCA 36


RESULT 3
US-08-219-633-25/c
; Sequence 25, Application US/08219633
; Patent No. 5599666
;  GENERAL INFORMATION:
;    APPLICANT:  Schumm, James W.
;    APPLICANT:  Puers, Christoph
;    TITLE OF INVENTION:  ALLELIC LADDERS FOR SHORT TANDEM REPEAT
;    TITLE OF INVENTION:  LOCI
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Ross & Stevens, S.C.
;      STREET:  One South Pinckney Street, P.O. Box 2599
;      CITY:  Madison
;      STATE:  Wisconsin
;      COUNTRY:  U.S.A.
;      ZIP:  53701-2599
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/219,633
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sara, Charles S.
;      REGISTRATION NUMBER:  30,492
;      REFERENCE/DOCKET NUMBER:  34506.019
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (608) 257-5353
;      TELEFAX:  (608) 257-9175
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  29 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-219-633-25

  Query Match             51.0%;  Score 14.8;  DB 1;  Length 29;
  Best Local Similarity   42.3%;  Pred. No. 6e+02;
  Matches   11;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          4 GAUUCUUUUUGUAAGCCCCAAGGGCU 29
              ||:: :: :: : | || | |||||:
Db         29 GATTATTCTTATCATCCACTAGGGCT 4


RESULT 4
US-08-515-236-25/c
; Sequence 25, Application US/08515236
; Patent No. 5674686
;  GENERAL INFORMATION:
;    APPLICANT:  Schumm, James W.
;    APPLICANT:  Puers, Christoph
;    TITLE OF INVENTION:  ALLELIC LADDERS FOR SHORT TANDEM REPEAT
;    TITLE OF INVENTION:  LOCI
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Ross & Stevens, S.C.
;      STREET:  One South Pinckney Street, P.O. Box 2599
;      CITY:  Madison
;      STATE:  Wisconsin
;      COUNTRY:  U.S.A.
;      ZIP:  53701-2599
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/515,236
;      FILING DATE:  15-AUG-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/219,633
;      FILING DATE:  28-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sara, Charles S.
;      REGISTRATION NUMBER:  30,492
;      REFERENCE/DOCKET NUMBER:  34506.019
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (608) 257-5353
;      TELEFAX:  (608) 257-9175
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  29 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-515-236-25

  Query Match             51.0%;  Score 14.8;  DB 1;  Length 29;
  Best Local Similarity   42.3%;  Pred. No. 6e+02;
  Matches   11;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          4 GAUUCUUUUUGUAAGCCCCAAGGGCU 29
```

```
              ||:: :: :: : | || | |||||:
Db          29 GATTATTCTTATCATCCACTAGGGCT 4


RESULT 5
US-08-761-950-25/c
; Sequence 25, Application US/08761950
; Patent No. 5783406
;  GENERAL INFORMATION:
;    APPLICANT:  Schumm, James W.
;    APPLICANT:  Puers, Christoph
;    TITLE OF INVENTION:  ALLELIC LADDERS FOR SHORT TANDEM REPEAT
;    TITLE OF INVENTION:  LOCI
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Ross & Stevens, S.C.
;      STREET:  One South Pinckney Street, P.O. Box 2599
;      CITY:  Madison
;      STATE:  Wisconsin
;      COUNTRY:  U.S.A.
;      ZIP:  53701-2599
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/761,950
;      FILING DATE:  09-DEC-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/515,236
;      FILING DATE:  15-AUG-1995
;      APPLICATION NUMBER:  US 08/219,633
;      FILING DATE:  28-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sara, Charles S.
;      REGISTRATION NUMBER:  30,492
;      REFERENCE/DOCKET NUMBER:  34506.019
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (608) 257-5353
;      TELEFAX:  (608) 257-9175
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  29 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-761-950-25

  Query Match             51.0%;  Score 14.8;  DB 1;  Length 29;
  Best Local Similarity   42.3%;  Pred. No. 6e+02;
  Matches   11;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          4 GAUUCUUUUUGUAAGCCCCAAGGGCU 29
              ||:: :: :: : | || | |||||:
Db          29 GATTATTCTTATCATCCACTAGGGCT 4


RESULT 6
US-08-632-575B-39/c
; Sequence 39, Application US/08632575B
; Patent No. 5843660
;  GENERAL INFORMATION:
;    APPLICANT:  Schumm, James W.
;    TITLE OF INVENTION:  Multiplex Amplification of
;    TITLE OF INVENTION:  Short Tandem Repeat Loci
;    NUMBER OF SEQUENCES:  61
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Promega Corporation
;      STREET:  2800 Woods Hollow Road
;      CITY:  Madison
;      STATE:  Wisconsin
;      COUNTRY:  U.S.A.
;      ZIP:  53711-5399
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette - 3.5 inch, 1.44 Mb
;      COMPUTER:  IBM compatible PC
;      OPERATING SYSTEM:  DOS, version 6.0
;      SOFTWARE:  WordPerfect 5.1 (DOS text format)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/632,575B
;      FILING DATE:  04/15/96
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/316,544
;      FILING DATE:  09/30/94
;  INFORMATION FOR SEQ ID NO:  39:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  29
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
;    POSITION IN GENOME:
;      MAP POSITION:  HUMvWFA31
US-08-632-575B-39

  Query Match             51.0%;  Score 14.8;  DB 2;  Length 29;
  Best Local Similarity   42.3%;  Pred. No. 6e+02;
  Matches   11;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          4 GAUUCUUUUUGUAAGCCCCAAGGGCU 29
              ||:: :: :: : | || | |||||:
Db          29 GATTATTCTTATCATCCACTAGGGCT 4


RESULT 7
US-09-327-229-31/c
; Sequence 31, Application US/09327229
; Patent No. 6221598
;  GENERAL INFORMATION:
;       APPLICANT: Schumm, James W.
;                  Sprecher, Cynthia J.
;                  Lins, Ann M.
;       TITLE OF INVENTION: MULTIPLEX AMPLIFICATION OF SHORT TANDEM
;                           REPEAT LOCI
;       NUMBER OF SEQUENCES: 32
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Ross & Stevens, S.C.
;            STREET: P. O. Box 2599
;            CITY: Madison
;            STATE: Wisconsin
;            COUNTRY: U.S.A.
;            ZIP: 53701-2599
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/327,229
;            FILING DATE: 07-Jun-1999
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/08/316,544
;            FILING DATE: 30-SEP-1994
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Sara, Charles S.
;            REGISTRATION NUMBER: 30,492
;            REFERENCE/DOCKET NUMBER: 34506.022
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 608-257-5353
;            TELEFAX: 608-257-9175
;  INFORMATION FOR SEQ ID NO: 31:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 29 base pairs
;            TYPE: nucleic acid
```

```
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;         MOLECULE TYPE: DNA (genomic)
;         SEQUENCE DESCRIPTION: SEQ ID NO: 31:
US-09-327-229-31

  Query Match             51.0%;  Score 14.8;  DB 3;  Length 29;
  Best Local Similarity   42.3%;  Pred. No. 6e+02;
  Matches   11;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          4 GAUUCUUUUUGUAAGCCCCAAGGGCU 29
              ||:: :: :: : | || | |||||:
Db         29 GATTATTCTTATCATCCACTAGGGCT 4


RESULT 8
US-09-199-542B-39/c
; Sequence 39, Application US/09199542B
; Patent No. 6479235
; GENERAL INFORMATION:
;  APPLICANT:  Schumm, James W.
;  APPLICANT:  Sprecher, Cynthia J.
;  TITLE OF INVENTION:  Multiplex Amplification of Short Tandem Repeat Loci
;  FILE REFERENCE:  16026/9212
;  CURRENT APPLICATION NUMBER: US/09/199,542B
;  CURRENT FILING DATE:  1998-11-25
;  PRIOR APPLICATION NUMBER:  US 08/316,544
;  PRIOR FILING DATE:  1994-09-30
;  PRIOR APPLICATION NUMBER:  US 08/632,575
;  PRIOR FILING DATE:  1996-04-15
;  NUMBER OF SEQ ID NOS:  110
;  SOFTWARE:  Word97 (converted to DOS text format)
; SEQ ID NO 39
;   LENGTH: 29
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   LOCATION: HUMvWFA31
US-09-199-542B-39

  Query Match             51.0%;  Score 14.8;  DB 4;  Length 29;
  Best Local Similarity   42.3%;  Pred. No. 6e+02;
  Matches   11;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          4 GAUUCUUUUUGUAAGCCCCAAGGGCU 29
              ||:: :: :: : | || | |||||:
Db         29 GATTATTCTTATCATCCACTAGGGCT 4


RESULT 9
PCT-US95-12608-31/c
; Sequence 31, Application PC/TUS9512608
;  GENERAL INFORMATION:
;    APPLICANT:  Schumm, James W.
;    APPLICANT:  Sprecher, Cynthia J.
;    APPLICANT:  Lins, Ann M.
;    TITLE OF INVENTION:  MULTIPLEX AMPLIFICATION OF SHORT TANDEM
;    TITLE OF INVENTION:  REPEAT LOCI
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Ross & Stevens, S.C.
;      STREET:  P. O. Box 2599
;      CITY:  Madison
;      STATE:  Wisconsin
;      COUNTRY:  U.S.A.
;      ZIP:  53701-2599
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/12608
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sara, Charles S.
;      REGISTRATION NUMBER:  30,492
;      REFERENCE/DOCKET NUMBER:  34506.022
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  608-257-5353
;      TELEFAX:  608-257-9175
;  INFORMATION FOR SEQ ID NO:  31:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  29 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
PCT-US95-12608-31

  Query Match             51.0%;  Score 14.8;  DB 5;  Length 29;
  Best Local Similarity   42.3%;  Pred. No. 6e+02;
  Matches   11;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          4 GAUUCUUUUUGUAAGCCCCAAGGGCU 29
              ||:: :: :: : | || | |||||:
Db         29 GATTATTCTTATCATCCACTAGGGCT 4


RESULT 10
US-08-632-575B-59/c
; Sequence 59, Application US/08632575B
; Patent No. 5843660
;  GENERAL INFORMATION:
;    APPLICANT:  Schumm, James W.
;    TITLE OF INVENTION:  Multiplex Amplification of
;    TITLE OF INVENTION:  Short Tandem Repeat Loci
;    NUMBER OF SEQUENCES:  61
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Promega Corporation
;      STREET:  2800 Woods Hollow Road
;      CITY:  Madison
;      STATE:  Wisconsin
;      COUNTRY:  U.S.A.
;      ZIP:  53711-5399
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette - 3.5 inch, 1.44 Mb
;      COMPUTER:  IBM compatible PC
;      OPERATING SYSTEM:  DOS, version 6.0
;      SOFTWARE:  WordPerfect 5.1 (DOS text format)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/632,575B
;      FILING DATE:  04/15/96
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/316,544
;      FILING DATE:  09/30/94
;  INFORMATION FOR SEQ ID NO:  59:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  32
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
;    POSITION IN GENOME:
;      MAP POSITION:  HUMvWFA31
US-08-632-575B-59

  Query Match             51.0%;  Score 14.8;  DB 2;  Length 32;
  Best Local Similarity   42.3%;  Pred. No. 6.1e+02;
  Matches   11;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          4 GAUUCUUUUUGUAAGCCCCAAGGGCU 29
              ||:: :: :: : | || | |||||:
Db         29 GATTATTCTTATCATCCACTAGGGCT 4
```

```
RESULT 11
US-09-199-542B-59/c
; Sequence 59, Application US/09199542B
; Patent No. 6479235
; GENERAL INFORMATION:
;  APPLICANT:  Schumm, James W.
;  APPLICANT:  Sprecher, Cynthia J.
;  TITLE OF INVENTION:  Multiplex Amplification of Short Tandem Repeat Loci
;  FILE REFERENCE:  16026/9212
;  CURRENT APPLICATION NUMBER: US/09/199,542B
;  CURRENT FILING DATE:  1998-11-25
;  PRIOR APPLICATION NUMBER:  US 08/316,544
;  PRIOR FILING DATE:  1994-09-30
;  PRIOR APPLICATION NUMBER:  US 08/632,575
;  PRIOR FILING DATE:  1996-04-15
;  NUMBER OF SEQ ID NOS:  110
;  SOFTWARE:  Word97 (converted to DOS text format)
; SEQ ID NO 59
;   LENGTH: 32
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   LOCATION: HUMvWFA31
US-09-199-542B-59

  Query Match             51.0%;  Score 14.8;  DB 4;  Length 32;
  Best Local Similarity   42.3%;  Pred. No. 6.1e+02;
  Matches   11;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          4 GAUUCUUUUUGUAAGCCCCAAGGGCU 29
              ||:: :: :: : | || | |||||:
Db         29 GATTATTCTTATCATCCACTAGGGCT 4


RESULT 12
US-09-063-733A-18
; Sequence 18, Application US/09063733A
; Patent No. 6372211
;  GENERAL INFORMATION:
;    APPLICANT:  Isaac, Barbara G.
;    APPLICANT:  Greenplate, John T.
;    APPLICANT:  Purcell, John P.
;    APPLICANT:  Romano, Charles P.
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR CONTROLLING
;    TITLE OF INVENTION:  INSECTS
;    NUMBER OF SEQUENCES:  58
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold White & Durkee
;      STREET:  PO Box 4433
;      CITY:  Houston
;      STATE:  TX
;      COUNTRY:  USA
;      ZIP:  77210-4433
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/063,733A
;      FILING DATE:  21-APR-1998
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Patterson, Melinda L.
;      REGISTRATION NUMBER:  33,062
;      REFERENCE/DOCKET NUMBER:  MOBT:022
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  713-787-1400
;      TELEFAX:  713-787-1440
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  25 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-063-733A-18

  Query Match             50.3%;  Score 14.6;  DB 4;  Length 25;
  Best Local Similarity   47.6%;  Pred. No. 7.2e+02;
  Matches   10;  Conservative    7;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 AAGAUUCUUUUUGUAAGCCCC 22
              ||| ::|: :::|:||  |||
Db          5 AAGCTTCTCTTTGTAATACCC 25


RESULT 13
US-08-486-969-46/c
; Sequence 46, Application US/08486969
; Patent No. 5843456
;  GENERAL INFORMATION:
;    APPLICANT:  Paoletti, Enzo
;    APPLICANT:  Maki, Joanne
;    TITLE OF INVENTION:  RECOMBINANT POXVIRUS - RABIES
;    TITLE OF INVENTION:  COMPOSITIONS AND COMBINATION COMPOSITIONS AND USES
;    NUMBER OF SEQUENCES:  55
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford, P.C.
;      STREET:  530 Fifth Avenue, 25th Floor
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  United States of America
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/486,969
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Frommer, William S.
;      REGISTRATION NUMBER:  25,506
;      REFERENCE/DOCKET NUMBER:  454310-2600
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;  INFORMATION FOR SEQ ID NO:  46:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  53 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-486-969-46

  Query Match             50.3%;  Score 14.6;  DB 2;  Length 53;
  Best Local Similarity   47.6%;  Pred. No. 8.3e+02;
  Matches   10;  Conservative    7;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          9 UUUUUGUAAGCCCCAAGGGCU 29
              :::::|:|||| | ||||:
Db         29 TTTTTGTAAGCTTCCCGGGCT 9


RESULT 14
US-09-827-998-1098/c
; Sequence 1098, Application US/09827998
; Patent No. 6656700
; GENERAL INFORMATION:
;  APPLICANT: Gu, Yizhong
;  APPLICANT:  Shannon, Mark
;  TITLE OF INVENTION: NOVEL ISOFORMS OF HUMAN PREGNANCY-ASSOCIATED PROTEIN E
;  FILE REFERENCE: MDhMORF-8
;  CURRENT APPLICATION NUMBER: US/09/827,998
```

```
;  CURRENT FILING DATE:  2001-04-06
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  NUMBER OF SEQ ID NOS: 1881
;  SOFTWARE: Aeomica Sequence Listing Engine
; Patent No. 6656700
; SEQ ID NO 1098
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-827-998-1098

  Query Match             49.0%;  Score 14.2;  DB 4;  Length 25;
  Best Local Similarity   47.4%;  Pred. No. 1.1e+03;
  Matches    9;  Conservative    7;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 UCUUUUUGUAAGCCCCAAG 25
              :|:::::|:|  ||| |||
Db         25 TCTTTTTGTAGTCCCTAAG 7


RESULT 15
US-09-827-998-1099/c
; Sequence 1099, Application US/09827998
; Patent No. 6656700
; GENERAL INFORMATION:
;  APPLICANT: Gu, Yizhong
;  APPLICANT:  Shannon, Mark
;  TITLE OF INVENTION: NOVEL ISOFORMS OF HUMAN PREGNANCY-ASSOCIATED PROTEIN E
;  FILE REFERENCE: MDhMORF-8
;  CURRENT APPLICATION NUMBER: US/09/827,998
;  CURRENT FILING DATE:  2001-04-06
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  NUMBER OF SEQ ID NOS: 1881
;  SOFTWARE: Aeomica Sequence Listing Engine
; Patent No. 6656700
; SEQ ID NO 1099
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-827-998-1099

  Query Match             49.0%;  Score 14.2;  DB 4;  Length 25;
  Best Local Similarity   47.4%;  Pred. No. 1.1e+03;
  Matches    9;  Conservative    7;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 UCUUUUUGUAAGCCCCAAG 25
              :|:::::|:|  ||| |||
Db         24 TCTTTTTGTAGTCCCTAAG 6


Search completed: March 23, 2004, 17:20:37
Job time : 58.3333 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 14:06:04 ; Search time 237.667 Seconds
(without alignments)
451.369 Million cell updates/sec

Title: US-09-310-844C-25
Perfect score: 29
Sequence: 1 aaagauucuuuuuguaagccccaagggcu 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2438257 seqs, 1849576744 residues

Total number of hits satisfying chosen parameters: 1514106

Minimum DB seq length: 0
Maximum DB seq length: 70

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		15.4	53.1	53	9	US-09-983-965-4754	Sequence 4754, Ap
2		15.2	52.4	47	15	US-10-349-143-1097	Sequence 1097, Ap
c 3		15.2	52.4	60	10	US-09-908-975-18725	Sequence 18725, A
4		15	51.7	39	14	US-10-116-519-18	Sequence 18, Appl
c 5		15	51.7	60	10	US-09-908-975-12187	Sequence 12187, A
6		15	51.7	65	10	US-09-908-975-4580	Sequence 4580, Ap
c 7		14.8	51.0	29	10	US-09-839-478-31	Sequence 31, Appl
8		14.6	50.3	25	14	US-10-005-530-18	Sequence 18, Appl
9		14.6	50.3	60	10	US-09-908-975-8435	Sequence 8435, Ap
10		14.6	50.3	60	10	US-09-908-975-18114	Sequence 18114, A
11		14.6	50.3	60	15	US-10-378-094-45	Sequence 45, Appl
12		14.6	50.3	60	16	US-10-231-494-24	Sequence 24, Appl
c 13		14.4	49.7	31	10	US-09-848-754A-6937	Sequence 6937, Ap
c 14		14.4	49.7	31	10	US-09-740-332-5660	Sequence 5660, Ap
c 15		14.4	49.7	31	10	US-09-817-879-5660	Sequence 5660, Ap
c 16		14.4	49.7	65	10	US-09-908-975-3924	Sequence 3924, Ap
17		14.4	49.7	65	10	US-09-908-975-30555	Sequence 30555, A
c 18		14.2	49.0	25	9	US-09-827-998-1098	Sequence 1098, Ap
c 19		14.2	49.0	25	9	US-09-827-998-1099	Sequence 1099, Ap
c 20		14.2	49.0	25	9	US-09-827-998-1100	Sequence 1100, Ap
c 21		14.2	49.0	25	9	US-09-827-998-1101	Sequence 1101, Ap
c 22		14.2	49.0	25	9	US-09-827-998-1102	Sequence 1102, Ap
c 23		14.2	49.0	25	9	US-09-827-998-1103	Sequence 1103, Ap
c 24		14.2	49.0	25	9	US-09-827-998-1104	Sequence 1104, Ap
25		14.2	49.0	25	14	US-10-098-263B-76444	Sequence 76444, A
26		14.2	49.0	26	14	US-10-287-919-545	Sequence 545, App
c 27		14.2	49.0	31	10	US-09-848-754A-6794	Sequence 6794, Ap
c 28		14.2	49.0	31	10	US-09-740-332-7606	Sequence 7606, Ap
c 29		14.2	49.0	31	10	US-09-817-879-7606	Sequence 7606, Ap
30		14.2	49.0	47	15	US-10-349-143-96	Sequence 96, Appl
c 31		14.2	49.0	50	15	US-10-131-827-2080	Sequence 2080, Ap
32		14.2	49.0	60	10	US-09-908-975-13840	Sequence 13840, A
c 33		14.2	49.0	60	10	US-09-908-975-13940	Sequence 13940, A
c 34		14.2	49.0	60	10	US-09-908-975-16808	Sequence 16808, A
35		14.2	49.0	65	9	US-09-783-590-1151	Sequence 1151, Ap
36		14.2	49.0	65	14	US-10-032-585-500	Sequence 500, App
37		14	48.3	41	10	US-09-852-385-2	Sequence 2, Appli
38		14	48.3	57	10	US-09-860-738C-46	Sequence 46, Appl
c 39		13.8	47.6	27	9	US-09-823-936-33	Sequence 33, Appl
c 40		13.8	47.6	27	13	US-10-193-142-10	Sequence 10, Appl
c 41		13.8	47.6	27	14	US-10-134-493-10	Sequence 10, Appl
c 42		13.8	47.6	27	14	US-10-061-216-10	Sequence 10, Appl
c 43		13.8	47.6	27	14	US-10-286-140-33	Sequence 33, Appl
c 44		13.8	47.6	31	10	US-09-740-332-9154	Sequence 9154, Ap
c 45		13.8	47.6	31	10	US-09-817-879-9154	Sequence 9154, Ap

ALIGNMENTS

RESULT 1
US-09-983-965-4754
; Sequence 4754, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 37-21(10297)C
; CURRENT APPLICATION NUMBER: US/09/983,965
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 09/465,231
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,678
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 4754
; LENGTH: 53
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; OTHER INFORMATION: Clone ID: 18-LIB34-011-Q1-E1-E5
US-09-983-965-4754

Query Match 53.1%; Score 15.4; DB 9; Length 53;
Best Local Similarity 44.0%; Pred. No. 3e+03;
Matches 11; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

```
Qy          5 AUUCUUUUUGUAAGCCCCAAGGGCU 29
              |::|::: :|:  |||   |||||:
Db         28 ATTCTTTGTGTTTGCCTTCAGGGCT 52
```

RESULT 2
US-10-349-143-1097

```
; Sequence 1097, Application US/10349143
; Publication No. US20040005584A1
; GENERAL INFORMATION:
;  APPLICANT: Cohen, Daniel
;  APPLICANT:  Blumenfeld, Marta
;  APPLICANT:  Chumakov, Ilya
;  TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
;  FILE REFERENCE: GENSET.020CP1
;  CURRENT APPLICATION NUMBER: US/10/349,143
;  CURRENT FILING DATE:  2003-01-21
;  PRIOR APPLICATION NUMBER: US/09/422,978
;  PRIOR FILING DATE: 1999-10-20
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 09/298,850
;  PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-21
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/109,732
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-23
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/082,614
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-21
;  NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 1097
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 24
;   OTHER INFORMATION: 99-2043-220 : polymorphic base A or T
US-10-349-143-1097

  Query Match             52.4%;  Score 15.2;  DB 15;  Length 47;
  Best Local Similarity   50.0%;  Pred. No. 3.6e+03;
  Matches   11;  Conservative    7;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          3 AGAUUCUUUUUGUAAGCCCCAA 24
              ||| :|:::::|:|| | || |
Db         15 AGACTCTTTWTGTAACCTCCCA 36


RESULT 3
US-09-908-975-18725/c
; Sequence 18725, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
;  APPLICANT: SHOSHAN, Avi
;  APPLICANT:  WASSERMAN, Alon
;  APPLICANT:  MINTZ, Eli
;  APPLICANT:  MINTZ, Liat
;  APPLICANT:  FAIGLER, Simchon
;  TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
;  TITLE OF INVENTION:  THAT POPULATE A TRANSCRIPTOME
;  FILE REFERENCE: 36688-0005
;  CURRENT APPLICATION NUMBER: US/09/908,975
;  CURRENT FILING DATE:  2001-07-20
;  PRIOR APPLICATION NUMBER: US 60/287,724
;  PRIOR FILING DATE: 2001-05-02
;  PRIOR APPLICATION NUMBER: US 60/221,607
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 32337
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18725
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-908-975-18725

  Query Match             52.4%;  Score 15.2;  DB 10;  Length 60;
  Best Local Similarity   60.0%;  Pred. No. 3.8e+03;
  Matches   12;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 UCUUUUUGUAAGCCCCAAGG 26
              :|::: :| |||||||| ||
Db         26 TCTTTCTGAAAGCCCCATGG 7


RESULT 4
US-10-116-519-18
; Sequence 18, Application US/10116519
; Publication No. US20030114373A1
; GENERAL INFORMATION:
;  APPLICANT: Bristol-Myers Squibb Company
;  TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL CYSTEINE PROTEASE OF THE CALPAIN
;  TITLE OF INVENTION:  SUPERFAMILY, CAN-12 AND VARIANTS THEREOF
;  FILE REFERENCE: D0124 NP
;  CURRENT APPLICATION NUMBER: US/10/116,519
;  CURRENT FILING DATE:  2002-04-03
;  PRIOR APPLICATION NUMBER: US 60/281,253
;  PRIOR FILING DATE: 2001-04-03
;  PRIOR APPLICATION NUMBER: US 60/288,768
;  PRIOR FILING DATE: 2001-05-04
;  PRIOR APPLICATION NUMBER: US 60/296,180
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: US 60/300,620
;  PRIOR FILING DATE: 2001-06-25
;  NUMBER OF SEQ ID NOS: 145
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
;   LENGTH: 39
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-116-519-18

  Query Match             51.7%;  Score 15;  DB 14;  Length 39;
  Best Local Similarity   56.5%;  Pred. No. 4.3e+03;
  Matches   13;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 GAUUCUUUUUGUAAGCCCCAAGG 26
              | ::| :: :| |||| ||||||
Db          5 GGTTCCTTCTGAAAGCTCCAAGG 27


RESULT 5
US-09-908-975-12187/c
; Sequence 12187, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
;  APPLICANT: SHOSHAN, Avi
;  APPLICANT:  WASSERMAN, Alon
;  APPLICANT:  MINTZ, Eli
;  APPLICANT:  MINTZ, Liat
;  APPLICANT:  FAIGLER, Simchon
;  TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE \
;  TITLE OF INVENTION:  THAT POPULATE A TRANSCRIPTOME
;  FILE REFERENCE: 36688-0005
;  CURRENT APPLICATION NUMBER: US/09/908,975
;  CURRENT FILING DATE:  2001-07-20
;  PRIOR APPLICATION NUMBER: US 60/287,724
;  PRIOR FILING DATE: 2001-05-02
;  PRIOR APPLICATION NUMBER: US 60/221,607
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 32337
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12187
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-908-975-12187

  Query Match             51.7%;  Score 15;  DB 10;  Length 60;
  Best Local Similarity   47.8%;  Pred. No. 4.6e+03;
  Matches   11;  Conservative    7;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          3 AGAUUCUUUUUGUAAGCCCCAAG 25
              |||::|::: :|:|  | | |||
Db         55 AGATTCTTTCTGTAGCCGCTAAG 33
```

```
RESULT 6
US-09-908-975-4580
; Sequence 4580, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
;  APPLICANT: SHOSHAN, Avi
;  APPLICANT:  WASSERMAN, Alon
;  APPLICANT:  MINTZ, Eli
;  APPLICANT:  MINTZ, Liat
;  APPLICANT:  FAIGLER, Simchon
;  TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
;  TITLE OF INVENTION:  THAT POPULATE A TRANSCRIPTOME
;  FILE REFERENCE: 36688-0005
;  CURRENT APPLICATION NUMBER: US/09/908,975
;  CURRENT FILING DATE:  2001-07-20
;  PRIOR APPLICATION NUMBER: US 60/287,724
;  PRIOR FILING DATE: 2001-05-02
;  PRIOR APPLICATION NUMBER: US 60/221,607
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 32337
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4580
;   LENGTH: 65
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
US-09-908-975-4580

  Query Match             51.7%;  Score 15;  DB 10;  Length 65;
  Best Local Similarity   56.5%;  Pred. No. 4.7e+03;
  Matches   13;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 AAGAUUCUUUUUGUAAGCCCCAA 24
              ||||: |:: ::| |||| |||
Db         34 AAGATGCTTCTTGAAAGCAACAA 56


RESULT 7
US-09-839-478-31/c
; Sequence 31, Application US/09839478
; Publication No. US20030180724A1
;   GENERAL INFORMATION:
;        APPLICANT: Schumm, James W.
;                   Sprecher, Cynthia J.
;                   Lins, Ann M.
;        TITLE OF INVENTION: MULTIPLEX AMPLIFICATION OF SHORT TANDEM
;                            REPEAT LOCI
;        NUMBER OF SEQUENCES: 32
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Ross & Stevens, S.C.
;             STREET: P. O. Box 2599
;             CITY: Madison
;             STATE: Wisconsin
;             COUNTRY: U.S.A.
;             ZIP: 53701-2599
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/839,478
;             FILING DATE: 20-Apr-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/316,544
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Sara, Charles S.
;             REGISTRATION NUMBER: 30,492
;             REFERENCE/DOCKET NUMBER: 34506.022
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 608-257-5353
;             TELEFAX: 608-257-9175
;   INFORMATION FOR SEQ ID NO: 31:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 29 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 31:
US-09-839-478-31

  Query Match             51.0%;  Score 14.8;  DB 10;  Length 29;
  Best Local Similarity   42.3%;  Pred. No. 4.9e+03;
  Matches   11;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          4 GAUUCUUUUUGUAAGCCCCAAGGGCU 29
              ||:: :: :: : | || | |||||:
Db         29 GATTATTCTTATCATCCACTAGGGCT 4


RESULT 8
US-10-005-530-18
; Sequence 18, Application US/10005530
; Publication No. US20030026795A1
; GENERAL INFORMATION:
;  APPLICANT: Isaac, Barbara G.
;  APPLICANT:  Greenplate, John T.
;  APPLICANT:  Purcell, John P.
;  APPLICANT:  Romano, Charles P.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING INSECTS
;  FILE REFERENCE: 11899.0022.DVUS01 (MOBT:022--2)
;  CURRENT APPLICATION NUMBER: US/10/005,530
;  CURRENT FILING DATE:  2001-10-26
;  PRIOR APPLICATION NUMBER: 09/063,733
;  PRIOR FILING DATE: 1998-04-21
;  PRIOR APPLICATION NUMBER: 60/044,504
;  PRIOR FILING DATE: 1997-04-21
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic Oligonucleotide
US-10-005-530-18

  Query Match             50.3%;  Score 14.6;  DB 14;  Length 25;
  Best Local Similarity   47.6%;  Pred. No. 5.9e+03;
  Matches   10;  Conservative    7;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 AAGAUUCUUUUUGUAAGCCCC 22
              ||| ::|: :::|:|| |||
Db          5 AAGCTTCTCTTTGTAATACCC 25


RESULT 9
US-09-908-975-8435
; Sequence 8435, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
;  APPLICANT: SHOSHAN, Avi
;  APPLICANT:  WASSERMAN, Alon
;  APPLICANT:  MINTZ, Eli
;  APPLICANT:  MINTZ, Liat
;  APPLICANT:  FAIGLER, Simchon
;  TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE \
;  TITLE OF INVENTION:  THAT POPULATE A TRANSCRIPTOME
;  FILE REFERENCE: 36688-0005
;  CURRENT APPLICATION NUMBER: US/09/908,975
;  CURRENT FILING DATE:  2001-07-20
;  PRIOR APPLICATION NUMBER: US 60/287,724
;  PRIOR FILING DATE: 2001-05-02
;  PRIOR APPLICATION NUMBER: US 60/221,607
```

; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8435
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-975-8435

Query Match 50.3%; Score 14.6; DB 10; Length 60;
Best Local Similarity 51.7%; Pred. No. 7.1e+03;
Matches 15; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

```
Qy          1 AAAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
              || ||  |:  ::|:|| ||||| |  |:
Db         12 AACGAAACTGATTGTAATCCCCAGGATCT 40
```

RESULT 10
US-09-908-975-18114
; Sequence 18114, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18114
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-975-18114

Query Match 50.3%; Score 14.6; DB 10; Length 60;
Best Local Similarity 52.4%; Pred. No. 7.1e+03;
Matches 11; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

```
Qy          9 UUUUUGUAAGCCCCAAGGGCU 29
              : ::: : || ||||||||||:
Db          7 TATTTCTGAGTCCCAAGGGCT 27
```

RESULT 11
US-10-378-094-45
; Sequence 45, Application US/10378094
; Publication No. US20030221201A1
; GENERAL INFORMATION:
; APPLICANT: PRIOR, Christopher P.
; APPLICANT: LAI, Char-Huei
; APPLICANT: SADEGHI, Homayoun
; APPLICANT: TURNER, Andrew
; TITLE OF INVENTION: MODIFIED TRANSFERRIN FUSION PROTEINS
; FILE REFERENCE: 54710-5001-01-US
; CURRENT APPLICATION NUMBER: US/10/378,094
; CURRENT FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: US 10/231,494
; PRIOR FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: US 60/334,059
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: US 60/315,745
; PRIOR FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 45
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic oligonucleotide encoding peptide with EPO activity
US-10-378-094-45

Query Match 50.3%; Score 14.6; DB 15; Length 60;
Best Local Similarity 52.4%; Pred. No. 7.1e+03;
Matches 11; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

```
Qy          6 UUCUUUUUGUAAGCCCCAAGG 26
              ::   :::|:||||| |||||
Db         36 TTGGGTTTGTAAGCCACAAGG 56
```

RESULT 12
US-10-231-494-24
; Sequence 24, Application US/10231494
; Publication No. US20040023334A1
; GENERAL INFORMATION:
; APPLICANT: Prior, Christopher P.
; TITLE OF INVENTION: Modified Transferrin Fusion Proteins
; FILE REFERENCE: 54710-5001-US
; CURRENT APPLICATION NUMBER: US/10/231,494
; CURRENT FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: US 60/315,745
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: US 60/334,059
; PRIOR FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: EPO mimetic
; OTHER INFORMATION: sequences
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(60)
US-10-231-494-24

Query Match 50.3%; Score 14.6; DB 16; Length 60;
Best Local Similarity 52.4%; Pred. No. 7.1e+03;
Matches 11; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

```
Qy          6 UUCUUUUUGUAAGCCCCAAGG 26
              ::   :::|:||||| |||||
Db         36 TTGGGTTTGTAAGCCACAAGG 56
```

RESULT 13
US-09-848-754A-6937/c
; Sequence 6937, Application US/09848754A
; Publication No. US20030073207A1
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; TITLE OF INVENTION: Enzymatic Nucleic Acid Treatment of Diseases or Conditions Relate
; TITLE OF INVENTION: Levels of Epidermal Growth Factor Receptors
; FILE REFERENCE: MBHB00-958-I (400/018)
; CURRENT APPLICATION NUMBER: US/09/848,754A
; CURRENT FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 9645
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6937
; LENGTH: 31
; TYPE: DNA
; ORGANISM: Artificial Sequence

```
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Enzymatic Nucleic acid
US-09-848-754A-6937

  Query Match             49.7%;  Score 14.4;  DB 10;  Length 31;
  Best Local Similarity   58.3%;  Pred. No. 7.6e+03;
  Matches   14;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 GAUUCUUUUUGUAAGCCCCAAGGG 27
              ||:  :: : |  ||||||||||||
Db         25 GATCGTTGTAGCTAGCCCCAAGGG 2


RESULT 14
US-09-740-332-5660/c
; Sequence 5660, Application US/09740332
; Publication No. US20030125270A1
; GENERAL INFORMATION:
;  APPLICANT: Ribozyme Pharmaceuticals Inc.
;  TITLE OF INVENTION: Enzymatic Nucleic Acid Treatment of Diseases or Conditions Relate
;  TITLE OF INVENTION:  Hepatitis C Virus Infection
;  FILE REFERENCE: RPI 400/003
;  CURRENT APPLICATION NUMBER: US/09/740,332
;  CURRENT FILING DATE:  2001-03-26
;  NUMBER OF SEQ ID NOS: 9704
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5660
;   LENGTH: 31
;   TYPE: DNA
;   ORGANISM: artificial sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION:
;   OTHER INFORMATION: DNAzyme
US-09-740-332-5660

  Query Match             49.7%;  Score 14.4;  DB 10;  Length 31;
  Best Local Similarity   58.3%;  Pred. No. 7.6e+03;
  Matches   14;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 AAGAUUCUUUUUGUAAGCCCCAAG 25
              ||||:  :: : |  |||||||||
Db         27 AAGATCGTTGTAGCTAGCCCCAAG 4


RESULT 15
US-09-817-879-5660/c
; Sequence 5660, Application US/09817879
; Publication No. US20030171311A1
; GENERAL INFORMATION:
;  APPLICANT: Ribozyme Pharmaceuticals Inc.
;  TITLE OF INVENTION: Enzymatic Nucleic Acid Treatment of Diseases or Conditions Relate
;  TITLE OF INVENTION:  Hepatitis C Virus Infection
;  FILE REFERENCE: MBHB00-801-F
;  CURRENT APPLICATION NUMBER: US/09/817,879
;  CURRENT FILING DATE:  2001-03-26
;  NUMBER OF SEQ ID NOS: 9703
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5660
;   LENGTH: 31
;   TYPE: DNA
;   ORGANISM: artificial sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION:
;   OTHER INFORMATION: DNAzyme
US-09-817-879-5660

  Query Match             49.7%;  Score 14.4;  DB 10;  Length 31;
  Best Local Similarity   58.3%;  Pred. No. 7.6e+03;
  Matches   14;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 AAGAUUCUUUUUGUAAGCCCCAAG 25
              ||||:  :: : |  |||||||||
Db         27 AAGATCGTTGTAGCTAGCCCCAAG 4


Search completed: March 23, 2004, 17:17:36
Job time : 240 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 14:02:34 ; Search time 1997.33 Seconds
(without alignments)
433.580 Million cell updates/sec

Title: US-09-310-844C-25
Perfect score: 29
Sequence: 1 aaagauucuuuuuguaagccccaagggcu 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 289680

Minimum DB seq length: 0
Maximum DB seq length: 70

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	18.4	63.4	70	9	AA516989	AA516989 vh89d02.r
c	2	17.4	60.0	67	9	AA708911	AA708911 zl64a10.s
c	3	16.6	57.2	70	9	AI609394	AI609394 tw93b03.x
c	4	16	55.2	51	12	BG361927	BG361927 gb49d10.y
c	5	15.8	54.5	58	9	AI824019	AI824019 wj29f03.x
	6	15.6	53.8	37	9	AI802260	AI802260 tj36g07.x
c	7	15.6	53.8	58	28	AZ834846	AZ834846 2M0117F18
	8	15.4	53.1	49	14	U44334	U44334 ENU44334 As
	9	15.2	52.4	61	9	AI318033	AI318033 ta75g02.x
	10	15.2	52.4	65	12	BM517546	BM517546 kj80g07.y
c	11	15	51.7	58	28	B02943	B02943 cSRL-163G2-
c	12	15	51.7	68	14	CD682098	CD682098 rj46c08.y
c	13	14.8	51.0	34	28	AZ840876	AZ840876 2M0138C08
	14	14.8	51.0	49	28	AZ576537	AZ576537 AST-T11C0
c	15	14.8	51.0	55	9	AI224478	AI224478 qx06d06.x
	16	14.8	51.0	64	10	BE636255	BE636255 SWOvAMCAQ
	17	14.8	51.0	67	29	CG588850	CG588850 OST240321
c	18	14.8	51.0	70	9	AL780467	AL780467 AL780467
	19	14.6	50.3	52	29	BX650715	BX650715 Arabidops
	20	14.6	50.3	53	29	AL940874	AL940874 Arabidops
	21	14.6	50.3	59	10	BE970792	BE970792 601680150
c	22	14.6	50.3	61	13	BQ479345	BQ479345 ku33d12.y
c	23	14.6	50.3	65	29	AL763793	AL763793 Arabidops
c	24	14.6	50.3	69	28	BZ768797	BZ768797 SALK_1407
c	25	14.6	50.3	70	28	BZ768791	BZ768791 SALK_1407
c	26	14.6	50.3	70	28	BZ768795	BZ768795 SALK_1407
	27	14.4	49.7	35	28	BH856246	BH856246 SALK_0811
	28	14.4	49.7	35	28	BH856247	BH856247 SALK_0811
c	29	14.4	49.7	37	28	AZ950243	AZ950243 2M0214C15
c	30	14.4	49.7	41	28	AZ598587	AZ598587 1M0413A04
	31	14.4	49.7	51	14	CF425249	CF425249 lad09g02.
	32	14.4	49.7	51	29	DME545740	AJ545740 Drosophil
	33	14.4	49.7	56	28	BZ665747	BZ665747 KG10262 D
c	34	14.4	49.7	57	12	BG362067	BG362067 gb47b08.y
c	35	14.4	49.7	58	9	AV953887	AV953887 AV953887
	36	14.4	49.7	64	9	AI321110	AI321110 d4c09nm.r
c	37	14.4	49.7	65	28	BH908271	BH908271 SALK_0468
c	38	14.4	49.7	66	12	BG361679	BG361679 gb48b04.y
c	39	14.4	49.7	66	29	CG485985	CG485985 OST20677
c	40	14.4	49.7	67	28	BH848343	BH848343 SALK_0678
c	41	14.4	49.7	67	29	CC517699	CC517699 CH240_364
c	42	14.4	49.7	67	29	CG474006	CG474006 OST1820 M
c	43	14.4	49.7	67	29	CG474744	CG474744 OST2811 M
c	44	14.4	49.7	67	29	CG475921	CG475921 OST4838 M
c	45	14.4	49.7	67	29	CG476132	CG476132 OST5408 M

ALIGNMENTS

RESULT 1
AA516989/c
LOCUS AA516989 70 bp mRNA linear EST 14-JUL-1997
DEFINITION vh89d02.r1 Knowles Solter mouse embryonic stem cell Mus musculus cDNA clone IMAGE:894147 5' similar to TR:G187568 G187568 MG44 ;, mRNA sequence.
ACCESSION AA516989
VERSION AA516989.1 GI:2256448
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 70)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:522107
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
High quality sequence stop: 1.

FEATURES Location/Qualifiers
source 1. .70
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J x DBA/2J F1"
/db_xref="taxon:10090"
/clone="IMAGE:894147"
/dev_stage="embryo"
/lab_host="DH10B"
/clone_lib="Knowles Solter mouse embryonic stem cell"
/note="Vector: pSPORT; Site_1: NotI; Site_2: SalI; Cloned unidirectionally from mRNA prepared from 800 blastocysts. Primer: SalI(dT): 5'-CGGTCGACCGTCGACCGTTTTTTTTTTTTTT-3'. cDNAs were cloned into the NotI/SalI sites of a pSPORT vector (Life Technologies). "

ORIGIN

Query Match 63.4%; Score 18.4; DB 9; Length 70;
Best Local Similarity 57.1%; Pred. No. 1.3e+04;
Matches 16; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 28
| |||::|:: :: || | ||||||||
Db 47 ACAGATTCTTCTTAGAAACACCAAGGGC 20

RESULT 2
AA708911/c
LOCUS AA708911 67 bp mRNA linear EST 24-DEC-1997
DEFINITION zl64a10.s1 Soares_pregnant_uterus_NbHPU Homo sapiens cDNA clone IMAGE:506682 3' similar to SW:RB32_HUMAN Q13637 RAS-RELATED PROTEIN RAB-32 ;, mRNA sequence.
ACCESSION AA708911
VERSION AA708911.1 GI:2718829
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 67)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.

FEATURES Location/Qualifiers
source 1. .67
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3812701"
/db_xref="taxon:9606"
/clone="IMAGE:506682"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Soares_pregnant_uterus_NbHPU"
/note="Organ: uterus; Vector: pT7T3-Pac; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' AACTGGAAGAATTCGCGGCCGCCTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by M. Fatima Bonaldo."

ORIGIN

Query Match 60.0%; Score 17.4; DB 9; Length 67;
Best Local Similarity 55.6%; Pred. No. 3e+04;
Matches 15; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy 3 AGAUUCUUUUUGUAAGCCCCAAGGGCU 29
||| |:: ::|:|| |||| | |||:
Db 51 AGAGACTTCTTGTAAACCCCCAAGGCT 25

RESULT 3
AI609394/c
LOCUS AI609394 70 bp mRNA linear EST 16-DEC-1999
DEFINITION tw93b03.x1 NCI_CGAP_HN6 Homo sapiens cDNA clone IMAGE:2267213 3' similar to SW:TCPH_HUMAN Q99832 T-COMPLEX PROTEIN 1, ETA SUBUNIT ;, mRNA sequence.
ACCESSION AI609394
VERSION AI609394.1 GI:4618561
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 70)
AUTHORS NCI/NIDR-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute / National Institute of Dental Research, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Chong Heon Lee, D.D.S., Mary May, J. Silvio Gutkind, Ph.D., Myung Hee Park, PhD.
cDNA Library Preparation: Stratagene, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Insert Length: 2028 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 1
POLYA=No.

FEATURES Location/Qualifiers
source 1. .70
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2267213"
/tissue_type="normal gingiva (cell line from immortalized keratinocytes)"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="NCI_CGAP_HN6"
/note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.3 kb. 5' adaptor sequence: 5' AATTCGGCACGAG 3' 3' GCCGTGCTC 5' 3' adaptor sequence: 5' (GA)10ACTAGTCTCGAGTTTTTTTTTTTTTTTT 3' EcoRI site appears to have been lost in a fraction of the

clones. Library constructed by Stratagene; available through Mary May, PhD (Oral and Pharyngeal Cancer Branch, National Institute of Dental and Craniofacial Research, NIH; mmay@yoda.nidr.nih.gov)."
ORIGIN

Query Match 57.2%; Score 16.6; DB 9; Length 70;
Best Local Similarity 47.8%; Pred. No. 5.9e+04;
Matches 11; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

```
Qy          6 UUCUUUUUGUAAGCCCCAAGGGC 28
              :: :::::|:  |||||||||  |
Db         54 TTTTTTTTGTGGGCCCCAAGGCC 32
```

RESULT 4
BG361927/c
LOCUS BG361927 51 bp mRNA linear EST 08-MAR-2001
DEFINITION gb49d10.y1 Moss EST library PPG Physcomitrella patens cDNA clone PEP_SOURCE_ID: 5', mRNA sequence.
ACCESSION BG361927
VERSION BG361927.1 GI:13251024
KEYWORDS EST.
SOURCE Physcomitrella patens
ORGANISM Physcomitrella patens
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
REFERENCE 1 (bases 1 to 51)
AUTHORS Quatrano,R., Bashiardes,S., Cove,D., Cuming,A., Knight,C., Clifton,S., Marra,M., Hillier,L., Pape,D., Martin,J., Wylie,T., Underwood,K., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Steptoe,M., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R., Waterston,R. and Wilson,R.
TITLE Leeds/Wash U Moss EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Ralph Quatrano
Leeds/Wash U Moss EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Libraries were constructed by Dr. Stavros Bashiardes as part of the Physcomitrella EST program (PEP) at the Univ. of Leeds (UK) and Washington Univ. in St. Louis (USA) DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Celia Knight (c.d.knight@leeds.ac.uk)
Seq primer: -40RP from Gibco.
FEATURES Location/Qualifiers
source 1. .51
/organism="Physcomitrella patens"
/mol_type="mRNA"
/db_xref="taxon:3218"
/clone="PEP_SOURCE_ID:"
/tissue_type="gametophore: 30 day old tissue, ammonium-grown"
/lab_host="DH10B"
/clone_lib="Moss EST library PPG"
/note="Vector: pAMP1; Construction of the cDNA library was performed by Dr. W. Gregg Clark using a modification of the cDNA synthesis protocol developed in the laboratory of Dr. Michael Lovett by Dr. Yulia Korshunova (personal communication). First polyA + RNA was isolated from total gametophore RNA using oligo dT magnetic beads. Following this, first strand cDNA synthesis was performed on the bead-bound polyA + RNA, during which an oligonucleotide anchor sequence was incorporated onto the 5'-ends of the cDNA. PCR amplification was then used to synthesize the second strand, to amplify the double stranded DNA, and to incorporate dUTP containing sequences into the ends of the double stranded cDNA. This DNA was size selected and cloned into pAMP1 using the CloneAMP pAMP1 System (Life Technologies, GibcoBRL) for cloning amplification products by a non-restriction site dependant process. The cloning was directional based on sequence asymmetry introduced at the ends during PCR amplification. The 3' cDNA ends are proximal to the NotI site of the multiple cloning site in pAMP1. This annealing mixture was transformed into chemically competent DH10B cells and selected for ampicillin resistant growth. The resulting clones (about 330,000) were pooled to make the library."
ORIGIN

Query Match 55.2%; Score 16; DB 12; Length 51;
Best Local Similarity 41.7%; Pred. No. 1.1e+05;
Matches 10; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

```
Qy          6 UUCUUUUUGUAAGCCCCAAGGGCU 29
              :: ::::: :||| ||||||  |:
Db         27 TTTTTTTTTTAAGACCCAAGAACT 4
```

RESULT 5
AI824019/c
LOCUS AI824019 58 bp mRNA linear EST 21-DEC-1999
DEFINITION wj29f03.x1 NCI_CGAP_Kid12 Homo sapiens cDNA clone IMAGE:2404253 3' similar to TR:O70278 O70278 MULTIPLE ENDOCRINE NEOPLASIA TYPE 1 CANDIDATE PROTEIN NUMBER 18. ;, mRNA sequence.
ACCESSION AI824019
VERSION AI824019.1 GI:5444690
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 58)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Insert Length: 806 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .58
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2404253"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Kid12"
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP_Kid5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1323912-1325831, 1471368-1472903 and 1492104-1493255). Subtraction by Bento Soares and M. Fatima Bonaldo. "
ORIGIN

```
  Query Match             54.5%;  Score 15.8;  DB 9;  Length 58;
  Best Local Similarity   44.4%;  Pred. No. 1.2e+05;
  Matches   12;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          3 AGAUUCUUUUUGUAAGCCCCAAGGGCU 29
              || :: ::::: ||| || ||| ||:
Db         56 AGCTTTTTTTTCCAAGTCCAAAGAGCT 30


RESULT 6
AI802260
LOCUS       AI802260                 37 bp    mRNA    linear   EST 13-DEC-1999
DEFINITION  tj36g07.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2143644 3'
            similar to TR:Q41120 Q41120 HYDROXYPROLINE-RICH GLYCOPROTEIN ;,
            mRNA sequence.
ACCESSION   AI802260
VERSION     AI802260.1  GI:5367732
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 37)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Life Technologies catalog #: 11548-013
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 1470   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .37
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2143644"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Pan1"
                     /note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.72 kb. Life Technologies catalog #:
                     11548-013"
ORIGIN

  Query Match             53.8%;  Score 15.6;  DB 9;  Length 37;
  Best Local Similarity   50.0%;  Pred. No. 1.6e+05;
  Matches   11;  Conservative    7;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AAAGAUUCUUUUUGUAAGCCCC 22
              |||  :: :::::| |||||||
Db          7 AAAATTTTTTTTTGGAAGCCCC 28


RESULT 7
AZ834846/c
LOCUS       AZ834846                 58 bp    DNA     linear   GSS 20-FEB-2001
DEFINITION  2M0117F18R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0117F18 R, genomic survey sequence.
ACCESSION   AZ834846
VERSION     AZ834846.1  GI:13004754
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0117  row: F  column: 18
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 58.
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0117F18"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

  Query Match             53.8%;  Score 15.6;  DB 28;  Length 58;
  Best Local Similarity   54.5%;  Pred. No. 1.4e+05;
  Matches   12;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 GAUUCUUUUUGUAAGCCCCAAG 25
              | ::|  :::|:|| |||||||
Db         24 GTTTCCCTTTGTAATCCCCAAG 3


RESULT 8
U44334
LOCUS       U44334                   49 bp    mRNA    linear   EST 03-APR-1996
DEFINITION  ENU44334 Aspergillus nidulans cleistothecium Emericella nidulans
            cDNA clone SE0762, mRNA sequence.
ACCESSION   U44334
VERSION     U44334.1  GI:1244997
KEYWORDS    EST.
SOURCE      Emericella nidulans (anamorph: Aspergillus nidulans)
  ORGANISM  Emericella nidulans
```

```
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
            Eurotiales; Trichocomaceae; Emericella.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   Lee,D., Lee,S., Hwang,H., Kim,J. and Chae,K.
  TITLE     Quantitative analysis of gene expression in sexual structures of
            Aspergillus nidulans by sequencing of 3'-directed cDNA clones
  JOURNAL   FEMS Microbiol. Lett. 138 (1), 71-76 (1996)
  MEDLINE   96236220
   PUBMED   8674973
COMMENT     Contact: Keon-Sang Chae
            Chonbuk National University
            Chonju, 561-756, S. Korea
            Tel: +82-652-70-3340
            Fax: +82-652-70-3345
            Email: chaeks@chonbuknms.chonbuk.ac.kr.
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="Emericella nidulans"
                     /mol_type="mRNA"
                     /strain="FGSC4"
                     /db_xref="taxon:162425"
                     /clone="SE0762"
                     /tissue_type="cleistothecium"
                     /cell_type="Hull cell"
                     /dev_stage="sexual"
                     /clone_lib="Aspergillus nidulans cleistothecium"
                     /note="3'-directed cDNA clones; single-pass sequencing"
ORIGIN

  Query Match             53.1%;  Score 15.4;  DB 14;  Length 49;
  Best Local Similarity   52.0%;  Pred. No. 1.8e+05;
  Matches   13;  Conservative    6;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          3 AGAUUCUUUUUGUAAGCCCCAAGGG 27
              |||::|:::   :||  ||||| ||
Db         20 AGATTCTTTCATTAACTCCCAACGG 44


RESULT 9
AI318033
LOCUS       AI318033                 61 bp    mRNA    linear   EST 16-FEB-1999
DEFINITION  ta75g02.x1 NCI_CGAP_HSC2 Homo sapiens cDNA clone IMAGE:2049938 3'
            similar to SW:RL34_HUMAN P49207 60S RIBOSOMAL PROTEIN L34. ;, mRNA
            sequence.
ACCESSION   AI318033
VERSION     AI318033.1  GI:4033793
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 61)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Herbert Morse, M.D., Michael R. Emmert-Buck,
            M.D., Ph.D.
             cDNA Library Preparation: David B. Krizman, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 384   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .61
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2049938"
                     /tissue_type="stem cell 34+/38+"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_HSC2"
                     /note="Organ: bone marrow; Vector: pAMP1; mRNA made from
                     bone marrow, stem cells 34+/38+, cDNA made by oligo-dT
                     priming.  Directionally cloned.  Size-selected on agarose
                     gel, average insert size 400 bp.  Primary library,
                     non-amplified.  "
ORIGIN

  Query Match             52.4%;  Score 15.2;  DB 9;  Length 61;
  Best Local Similarity   53.6%;  Pred. No. 2e+05;
  Matches   15;  Conservative    5;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 AAAGAUUCUUUUUGUAAGCCCCAAGGGC 28
              || | ::| :  :|:| | || ||||||
Db          1 AAGGGTTCGTGCTGTATGACCTAAGGGC 28


RESULT 10
BM517546
LOCUS       BM517546                 65 bp    mRNA    linear   EST 15-FEB-2002
DEFINITION  kj80g07.y1 Ascaris suum female head SL1 TOPO v1 Murphy Chiapelli
            McCarter Ascaris suum cDNA 5', mRNA sequence.
ACCESSION   BM517546
VERSION     BM517546.1  GI:18688698
KEYWORDS    EST.
SOURCE      Ascaris suum (pig roundworm)
  ORGANISM  Ascaris suum
            Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida;
            Ascaridoidea; Ascarididae; Ascaris.
REFERENCE   1  (bases 1 to 65)
  AUTHORS   McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J.,
            Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B.,
            Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C.,
            Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
            Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T.,
            Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
            McCann,R., Waterston,R. and Wilson,R.
  TITLE     The Washington Univ. Nematode EST Project, 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: McCarter JP
            The Washington Univ. Nematode EST Project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            The library was constructed by Claire Murphy, Brandi Chiapelli, and
            Dr. James McCarter at Washington University, St. Louis. DNA
            Sequencing by: Washington University Genome Sequencing Center.
FEATURES             Location/Qualifiers
     source          1. .65
                     /organism="Ascaris suum"
                     /mol_type="mRNA"
                     /db_xref="taxon:6253"
                     /sex="Female"
                     /tissue_type="Head"
                     /dev_stage="Adult"
                     /lab_host="DH10B"
                     /clone_lib="Ascaris suum female head SL1 TOPO v1 Murphy
                     Chiapelli McCarter"
                     /note="Vector: pCRII-TOPO (Invitrogen); Site_1: EcoRI;
                     Site_2: EcoRI; The library was constructed by Claire
                     Murphy, Brandi Chiapelli,  and Dr. James McCarter at
                     Washington University, St. Louis. Oligo(dT)-SL1  PCR based
                     library. Ascaris suum female head cDNA PCR products of
                     size  >400 nucleotides containing SL1 on the 5' end and
```

oligo(dT) on the 3' end were non-directionally cloned into pCRII-TOPO(Invitrogen) following the Topo TA cloning protocol. Dissected nematode tissues were provided by Dr. Alan Scott (ascott@jhsph.edu) of the School of Public Hygiene and Public Health at John Hopkins University in Baltimore, MD"

ORIGIN

```
  Query Match             52.4%;  Score 15.2;  DB 12;  Length 65;
  Best Local Similarity   42.9%;  Pred. No. 2e+05;
  Matches   12;  Conservative    8;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 AAGAUUCUUUUUGUAAGCCCCAAGGGCU 29
              ||| ::|: ::: : || || |||  |:
Db          1 AAGGTTCTGTTTATGAGACCGAAGATCT 28
```

RESULT 11
B02943/c
LOCUS B02943 58 bp DNA linear GSS 13-JUL-1996
DEFINITION cSRL-163G2-u cSRL flow sorted Chromosome 11 specific cosmid Homo sapiens genomic clone cSRL-163G2, genomic survey sequence.
ACCESSION B02943
VERSION B02943.1 GI:1412221
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 58)
AUTHORS Evans,G.A., Burbee,D., Davies,C., Hahner,L., Oliver,T., Gilbert,M., Jones,D., Ward,T., Gillilan,E., Schagemann,J., Probst,S. ., Harris,J., DeFord,J., McFarland,J., Burzinski,K., Khan,M., Kupfer,K. and Garner,H.R.
TITLE Genomic Sequence Sampled Map of Chromosome 11
JOURNAL Unpublished (1996)
COMMENT Contact: Evans GA, Shane Probst
McDermott Center for Human Growth and Development
University of Texas Southwestern Medical Center At Dallas
5323 Harry Hines Blvd, Dallas TX 75235-8591
Tel: 214-648-1600
Fax: 214-648-1666
Email: gevans@utsw.swmed.edu, shane@mcdermott.swmed.edu
Seq primer: T7
Class: cosmid ends
High quality sequence stop: 58.
FEATURES Location/Qualifiers
source 1. .58
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="cSRL-163G2"
/sex="female"
/cell_type="chimeric hamster somatic cell hybrid"
/clone_lib="cSRL flow sorted Chromosome 11 specific cosmid"
/note="Vector: sCos-1; Human Chromosome 11 specific cosmid library prepared from flow sorted human Chromosome 11 derived from Chinese Hampster Ovary (CHO) monochromosomal somatic cell hybrid, J1"
ORIGIN

```
  Query Match             51.7%;  Score 15;  DB 28;  Length 58;
  Best Local Similarity   43.5%;  Pred. No. 2.4e+05;
  Matches   10;  Conservative    8;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          6 UUCUUUUUGUAAGCCCCAAGGGC 28
              :: ::::: : | ||| ||||||
Db         41 TTTTTTTTTGATCCCAAAGGGC 19
```

RESULT 12
CD682098/c
LOCUS CD682098 68 bp mRNA linear EST 24-JUN-2003
DEFINITION rj46c08.y1 Meloidogyne chitwoodi egg SL1 TOPO v1 Meloidogyne chitwoodi cDNA 5', mRNA sequence.
ACCESSION CD682098
VERSION CD682098.1 GI:32183864
KEYWORDS EST.
SOURCE Meloidogyne chitwoodi
ORGANISM Meloidogyne chitwoodi
Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina; Tylenchoidea; Heteroderidae; Meloidogyninae; Meloidogyne.
REFERENCE 1 (bases 1 to 68)
AUTHORS McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE The Washington Univ. Nematode EST Project, 1999
JOURNAL Unpublished (1999)
COMMENT Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
The library was constructed by Claire Murphy and Dr. Makedonka Dautova at Washington University, St. Louis (mdautova@watson.wustl.edu). Oligo(dT)-SL1 PCR based library. Meloidogyne chitwoodi egg cDNA PCR products of size >400 nucleotides containing SL1 on the 5' end and oligo(dT) on the 3' end were non-direction ally cloned into pCRII-TOPO(Invitrogen) following the Topo TA cloning protocol. Eggs were provided by Dr. David Bird of North Carolina State University, Raleigh, NC (david_bird@ncsu.edu).
Putative full length read
The vector to vector length is 69
Seq primer: SL1 primer.
FEATURES Location/Qualifiers
source 1. .68
/organism="Meloidogyne chitwoodi"
/mol_type="mRNA"
/db_xref="taxon:59747"
/dev_stage="eggs"
/lab_host="DH10B"
/clone_lib="Meloidogyne chitwoodi egg SL1 TOPO v1"
/note="Vector: pCRII-TOPO (Invitrogen); Site_1: EcoRI; Site_2: EcoRI; The library was constructed by Claire Murphy and Dr. Makedonka Dautova at Washington University, St. Louis. Oligo(dT)-SL1 PCR based library. Meloidogyne chitwoodi egg cDNA PCR products of size >400 nucleotides containing SL1 on the 5' end and oligo(dT) on the 3' end were non-directionally cloned into pCRII-TOPO(Invitrogen) following the Topo TA cloning protocol. Eggs were provided by Dr. David Bird of North Carolina State University, Raleigh, NC (david_bird@ncsu.edu)."
ORIGIN

```
  Query Match             51.7%;  Score 15;  DB 14;  Length 68;
  Best Local Similarity   47.8%;  Pred. No. 2.4e+05;
  Matches   11;  Conservative    7;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          6 UUCUUUUUGUAAGCCCCAAGGGC 28
              ::|: ::: :||||| ||  |||
Db         57 TTCTCTTTTTAAGCCTCACCGGC 35
```

RESULT 13
AZ840876/c
LOCUS AZ840876 34 bp DNA linear GSS 20-FEB-2001
DEFINITION 2M0138C08R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0138C08 R, genomic survey sequence.

```
ACCESSION   AZ840876
VERSION     AZ840876.1  GI:13010784
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 34)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000  Std Error: 0.00
            Plate: 0138  row: C  column: 08
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 34.
FEATURES             Location/Qualifiers
     source          1. .34
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0138C08"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

  Query Match             51.0%;  Score 14.8;  DB 28;  Length 34;
  Best Local Similarity   57.7%;  Pred. No. 3.2e+05;
  Matches   15;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          3 AGAUUCUUUUUGUAAGCCCCAAGGGC 28
              | |:  : :: | |||| ||||||||
Db         26 ATATAATCTTCGAAAGCACCAAGGGC 1


RESULT 14
AZ576537
LOCUS       AZ576537                 49 bp    DNA     linear   GSS 06-DEC-2000
DEFINITION  AST-T11C0260\( Genetrap T47D Human Breast Carcinoma Library Homo
            sapiens genomic 5', genomic survey sequence.
ACCESSION   AZ576537
VERSION     AZ576537.1  GI:11562848
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   Henkel,G., Liyanage,M., Pratt,E., Huang,D., Riley,M.,
            Bernardino,A., Durick,K. and Pollok,B.
  TITLE     Exon-trap tags from a T47D GenomeScreen(TM) Library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Greg Henkel
            Gene Expression
            Aurora Biosciences Corp.
            11010 Torreyana Road, San Diego, CA 92121, USA
            Tel: 8584048436
            Fax: 8584046719
            Email: henkelg@aurorabio.com
            Pools of cells were isolated from a GenomeScreen(TM) library. The
            library of cells was generated by retroviral integration of a gene
            tagging element consisting of: 1) A promoterless beta-lactamase
            proceeded by a splice acceptor as a reporter for gene expression;
            2) A promoter driving neomycin resistance followed by a splice
            donor to trap downstream exons. 3' RACE from neomycin gene was
            performed using total RNA from isolated pools. Output was shotgun
            cloned in pAmp-1 and used to transform DH5-alpha competent
            bacteria. 5' ends of reported sequences were immediately preceded
            by splice donor from the trapping co nstruct.
            Class: exon-trapped.
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /tissue_type="Carcinoma"
                     /cell_type="Epithelial"
                     /cell_line="T47D"
                     /clone_lib="Genetrap T47D Human Breast Carcinoma Library"
                     /note="Organ: Breast; Vector: pAmp-1; 3' RACE of total RNA
                     from genetrap pools; shotgun clone in pAmp-1 and  used to
                     transform DH5-alpha competent bacteria."
ORIGIN

  Query Match             51.0%;  Score 14.8;  DB 28;  Length 49;
  Best Local Similarity   57.7%;  Pred. No. 3e+05;
  Matches   15;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 AAAGAUUCUUUUUGUAAGCCCCAAGG 26
              |||||  |::  :|:||||  ||| |
Db          6 AAAGAGACTTCCTGTAAGCGGCAAAG 31


RESULT 15
AI224478/c
LOCUS       AI224478                 55 bp    mRNA    linear   EST 21-DEC-1998
DEFINITION  qx06d06.x1 NCI_CGAP_Lym12 Homo sapiens cDNA clone IMAGE:2000555 3'
            similar to TR:O23462 O23462 HYPOTHETICAL 18.0 KD PROTEIN. ;, mRNA
            sequence.
ACCESSION   AI224478
VERSION     AI224478.1  GI:3807191
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 55)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
```

```
           unknown library type
           Trace considered overall poor quality
           Insert Length: 1214   Std Error: 0.00
           Seq primer: -40UP from Gibco
           High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .55
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2000555"
                     /tissue_type="lymphoma, follicular mixed small and large
                     cell"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Lym12"
                     /note="Organ: lymph node; Vector: pCMV-SPORT6; Site_1:
                     SalI; Site_2: NotI; Cloned unidirectionally.  Primer:
                     Oligo dT.  Average insert size 1.25 kb. Life Technologies
                     catalog #: 11547-015"
ORIGIN

  Query Match             51.0%;  Score 14.8;  DB 9;  Length 55;
  Best Local Similarity   38.5%;  Pred. No. 2.9e+05;
  Matches   10;  Conservative    9;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          4 GAUUCUUUUUGUAAGCCCCAAGGGCU 29
              | :: ::::: :   ||||||||| :
Db         30 GTTTTTTTTTTTTCCCCCAAGGGTT 5


Search completed: March 23, 2004, 17:06:02
Job time : 2007.33 secs
```